

PCTWORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶ : C07H 21/00, C12N 1/15, 1/21, 5/10, 15/11, 15/63		A1	(11) International Publication Number: WO 99/31117 (43) International Publication Date: 24 June 1999 (24.06.99)
(21) International Application Number: PCT/US98/27059 (22) International Filing Date: 17 December 1998 (17.12.98)		<p>20874 (US). RUBEN, Steven, M. [US/US]; 18528 Heritage Hills Drive, Olney, MD 20832 (US). CARTER, Kenneth, C. [US/US]; 11601 Brandy Hall Lane, North Potomac, MD 20878 (US). SHI, Yang-gu [CN/US]; 437 West Side Drive, Gaithersburg, MD 20878 (US). ROSEN, Craig, A. [US/US]; 22400 Rolling Hill Road, Laytonsville, MD 20882 (US). SOPPET, Daniel, R. [US/US]; 15050 Stillfield Place, Centreville, VA 22020 (US). KYAW, Hla [MM/US]; 520 Sugarbush Circle, Frederick, MD 21703 (US). WEI, Ying-Fei [CN/US]; 1714-C Marina Court, San Mateo, CA 94403 (US). FLORENCE, Kimberly [US/US]; 12805 Atlantic Avenue, Rockville, MD 20851 (US). DUAN, Roxanne, D. [US/US]; 5515 Northfield Road, Bethesda, MD 20817 (US). FLORENCE, Charles [US/US]; 12805 Atlantic Avenue, Rockville, MD 20851 (US). GREENE, John, M. [US/US]; 872 Diamond Drive, Gaithersburg, MD 20878 (US). FENG, Ping [CN/US]; 4 Relda Court, Gaithersburg, MD 20878 (US). FERRIE, Ann, M. [US/US]; 120 Fox Run Drive, Tewksbury, MA 01876 (US). YU, Guo-Liang [CN/US]; 1714-C Marina Court, San Mateo, CA 94403 (US). JANAT, Fouad [SY/US]; 140 High Street, No. 202, Westerly, RI 02891 (US). NI, Jian [CN/US]; 5502 Manorfield Road, Rockville, MD 20853 (US).</p> <p>(74) Agents: BROOKES, A., Anders et al.; Human Genome Sciences, Inc., 9410 Key West Avenue, Rockville, MD 20850 (US).</p> <p>(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, HR, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).</p> <p>Published With international search report.</p>	
(30) Priority Data: 60/070,923 18 December 1997 (18.12.97) US 60/068,007 18 December 1997 (18.12.97) US 60/068,057 18 December 1997 (18.12.97) US 60/068,006 18 December 1997 (18.12.97) US 60/068,008 18 December 1997 (18.12.97) US 60/068,054 18 December 1997 (18.12.97) US 60/068,064 18 December 1997 (18.12.97) US 60/068,053 18 December 1997 (18.12.97) US 60/068,169 19 December 1997 (19.12.97) US 60/068,368 19 December 1997 (19.12.97) US 60/068,367 19 December 1997 (19.12.97) US 60/068,369 19 December 1997 (19.12.97) US 60/068,365 19 December 1997 (19.12.97) US			
(71) Applicant (for all designated States except US): HUMAN GENOME SCIENCES, INC. [US/US]; 9410 Key West Avenue, Rockville, MD 20850 (US). (72) Inventors; and (75) Inventors/Applicants (for US only): MOORE, Paul, A. [US/US]; 19005 Leatherbark Drive, Germantown, MD			
(54) Title: 110 HUMAN SECRETED PROTEINS			
(57) Abstract <p>The present invention relates to novel human secreted proteins and isolated nucleic acids containing the coding regions of the genes encoding such proteins. Also provided are vectors, host cells, antibodies, and recombinant methods for producing human secreted proteins. The invention further relates to diagnostic and therapeutic methods useful for diagnosing and treating disorders related to these novel human secreted proteins.</p>			

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

110 Human Secreted Proteins

Field of the Invention

This invention relates to newly identified polynucleotides and the polypeptides encoded by these polynucleotides, uses of such polynucleotides and polypeptides, and
5 their production.

Background of the Invention

Unlike bacterium, which exist as a single compartment surrounded by a membrane, human cells and other eucaryotes are subdivided by membranes into many functionally distinct compartments. Each membrane-bounded compartment, or
10 organelle, contains different proteins essential for the function of the organelle. The cell uses "sorting signals," which are amino acid motifs located within the protein, to target proteins to particular cellular organelles.

One type of sorting signal, called a signal sequence, a signal peptide, or a leader sequence, directs a class of proteins to an organelle called the endoplasmic reticulum
15 (ER). The ER separates the membrane-bounded proteins from all other types of proteins. Once localized to the ER, both groups of proteins can be further directed to another organelle called the Golgi apparatus. Here, the Golgi distributes the proteins to vesicles, including secretory vesicles, the cell membrane, lysosomes, and the other organelles.

20 Proteins targeted to the ER by a signal sequence can be released into the extracellular space as a secreted protein. For example, vesicles containing secreted proteins can fuse with the cell membrane and release their contents into the extracellular space - a process called exocytosis. Exocytosis can occur constitutively or after receipt of a triggering signal. In the latter case, the proteins are stored in secretory vesicles (or
25 secretory granules) until exocytosis is triggered. Similarly, proteins residing on the cell membrane can also be secreted into the extracellular space by proteolytic cleavage of a "linker" holding the protein to the membrane.

Despite the great progress made in recent years, only a small number of genes encoding human secreted proteins have been identified. These secreted proteins include
30 the commercially valuable human insulin, interferon, Factor VIII, human growth hormone, tissue plasminogen activator, and erythropoietin. Thus, in light of the pervasive role of secreted proteins in human physiology, a need exists for identifying and characterizing novel human secreted proteins and the genes that encode them. This knowledge will allow one to detect, to treat, and to prevent medical disorders by using
35 secreted proteins or the genes that encode them.

Summary of the Invention

The present invention relates to novel polynucleotides and the encoded polypeptides. Moreover, the present invention relates to vectors, host cells, antibodies, and recombinant methods for producing the polypeptides and polynucleotides. Also provided are diagnostic methods for detecting disorders related to the polypeptides, and therapeutic methods for treating such disorders. The invention further relates to screening methods for identifying binding partners of the polypeptides.

Detailed Description

Definitions

The following definitions are provided to facilitate understanding of certain terms used throughout this specification.

In the present invention, "isolated" refers to material removed from its original environment (e.g., the natural environment if it is naturally occurring), and thus is altered "by the hand of man" from its natural state. For example, an isolated polynucleotide could be part of a vector or a composition of matter, or could be contained within a cell, and still be "isolated" because that vector, composition of matter, or particular cell is not the original environment of the polynucleotide.

In the present invention, a "secreted" protein refers to those proteins capable of being directed to the ER, secretory vesicles, or the extracellular space as a result of a signal sequence, as well as those proteins released into the extracellular space without necessarily containing a signal sequence. If the secreted protein is released into the extracellular space, the secreted protein can undergo extracellular processing to produce a "mature" protein. Release into the extracellular space can occur by many mechanisms, including exocytosis and proteolytic cleavage.

As used herein, a "polynucleotide" refers to a molecule having a nucleic acid sequence contained in SEQ ID NO:X or the cDNA contained within the clone deposited with the ATCC. For example, the polynucleotide can contain the nucleotide sequence of the full length cDNA sequence, including the 5' and 3' untranslated sequences, the coding region, with or without the signal sequence, the secreted protein coding region, as well as fragments, epitopes, domains, and variants of the nucleic acid sequence. Moreover, as used herein, a "polypeptide" refers to a molecule having the translated amino acid sequence generated from the polynucleotide as broadly defined.

In the present invention, the full length sequence identified as SEQ ID NO:X was often generated by overlapping sequences contained in multiple clones (contig

analysis). A representative clone containing all or most of the sequence for SEQ ID NO:X was deposited with the American Type Culture Collection ("ATCC"). As shown in Table 1, each clone is identified by a cDNA Clone ID (Identifier) and the ATCC Deposit Number. The ATCC is located at 10801 University Boulevard,
5 Manassas, Virginia 20110-2209, USA. The ATCC deposit was made pursuant to the terms of the Budapest Treaty on the international recognition of the deposit of microorganisms for purposes of patent procedure.

A "polynucleotide" of the present invention also includes those polynucleotides capable of hybridizing, under stringent hybridization conditions, to sequences contained
10 in SEQ ID NO:X, the complement thereof, or the cDNA within the clone deposited with the ATCC. "Stringent hybridization conditions" refers to an overnight incubation at 42°C in a solution comprising 50% formamide, 5x SSC (750 mM NaCl, 75 mM sodium citrate), 50 mM sodium phosphate (pH 7.6), 5x Denhardt's solution, 10% dextran sulfate, and 20 µg/ml denatured, sheared salmon sperm DNA, followed by washing the
15 filters in 0.1x SSC at about 65°C.

Also contemplated are nucleic acid molecules that hybridize to the polynucleotides of the present invention at lower stringency hybridization conditions. Changes in the stringency of hybridization and signal detection are primarily accomplished through the manipulation of formamide concentration (lower percentages
20 of formamide result in lowered stringency); salt conditions, or temperature. For example, lower stringency conditions include an overnight incubation at 37°C in a solution comprising 6X SSPE (20X SSPE = 3M NaCl; 0.2M NaH₂PO₄; 0.02M EDTA, pH 7.4), 0.5% SDS, 30% formamide, 100 µg/ml salmon sperm blocking DNA; followed by washes at 50°C with 1XSSPE, 0.1% SDS. In addition, to achieve even
25 lower stringency, washes performed following stringent hybridization can be done at higher salt concentrations (e.g. 5X SSC).

Note that variations in the above conditions may be accomplished through the inclusion and/or substitution of alternate blocking reagents used to suppress background in hybridization experiments. Typical blocking reagents include
30 Denhardt's reagent, BLOTTO, heparin, denatured salmon sperm DNA, and commercially available proprietary formulations. The inclusion of specific blocking reagents may require modification of the hybridization conditions described above, due to problems with compatibility.

Of course, a polynucleotide which hybridizes only to polyA+ sequences (such
35 as any 3' terminal polyA+ tract of a cDNA shown in the sequence listing), or to a

complementary stretch of T (or U) residues, would not be included in the definition of "polynucleotide," since such a polynucleotide would hybridize to any nucleic acid molecule containing a poly (A) stretch or the complement thereof (e.g., practically any double-stranded cDNA clone).

5 The polynucleotide of the present invention can be composed of any polyribonucleotide or polydeoxribonucleotide, which may be unmodified RNA or DNA or modified RNA or DNA. For example, polynucleotides can be composed of single- and double-stranded DNA, DNA that is a mixture of single- and double-stranded regions, single- and double-stranded RNA, and RNA that is mixture of single- and double-stranded regions, hybrid molecules comprising DNA and RNA that may be
10 single-stranded or, more typically, double-stranded or a mixture of single- and double-stranded regions. In addition, the polynucleotide can be composed of triple-stranded regions comprising RNA or DNA or both RNA and DNA. A polynucleotide may also contain one or more modified bases or DNA or RNA backbones modified for stability
15 or for other reasons. "Modified" bases include, for example, tritylated bases and unusual bases such as inosine. A variety of modifications can be made to DNA and RNA; thus, "polynucleotide" embraces chemically, enzymatically, or metabolically modified forms.

 The polypeptide of the present invention can be composed of amino acids joined
20 to each other by peptide bonds or modified peptide bonds, i.e., peptide isosteres, and may contain amino acids other than the 20 gene-encoded amino acids. The polypeptides may be modified by either natural processes, such as posttranslational processing, or by chemical modification techniques which are well known in the art. Such modifications are well described in basic texts and in more detailed monographs,
25 as well as in a voluminous research literature. Modifications can occur anywhere in a polypeptide, including the peptide backbone, the amino acid side-chains and the amino or carboxyl termini. It will be appreciated that the same type of modification may be present in the same or varying degrees at several sites in a given polypeptide. Also, a given polypeptide may contain many types of modifications. Polypeptides may be
30 branched, for example, as a result of ubiquitination, and they may be cyclic, with or without branching. Cyclic, branched, and branched cyclic polypeptides may result from posttranslation natural processes or may be made by synthetic methods. Modifications include acetylation, acylation, ADP-ribosylation, amidation, covalent attachment of flavin, covalent attachment of a heme moiety, covalent attachment of a
35 nucleotide or nucleotide derivative, covalent attachment of a lipid or lipid derivative, covalent attachment of phosphatidylinositol, cross-linking, cyclization, disulfide bond formation, demethylation, formation of covalent cross-links, formation of cysteine,

formation of pyroglutamate, formylation, gamma-carboxylation, glycosylation, GPI anchor formation, hydroxylation, iodination, methylation, myristoylation, oxidation, pegylation, proteolytic processing, phosphorylation, prenylation, racemization, selenoylation, sulfation, transfer-RNA mediated addition of amino acids to proteins such as arginylation, and ubiquitination. (See, for instance, PROTEINS -
5 STRUCTURE AND MOLECULAR PROPERTIES, 2nd Ed., T. E. Creighton, W. H. Freeman and Company, New York (1993); POSTTRANSLATIONAL COVALENT MODIFICATION OF PROTEINS, B. C. Johnson, Ed., Academic Press, New York, pgs. 1-12 (1983); Seifter et al., Meth Enzymol 182:626-646 (1990);
10 Rattan et al., Ann NY Acad Sci 663:48-62 (1992).)

"SEQ ID NO:X" refers to a polynucleotide sequence while "SEQ ID NO:Y" refers to a polypeptide sequence, both sequences identified by an integer specified in Table 1.

"A polypeptide having biological activity" refers to polypeptides exhibiting
15 activity similar, but not necessarily identical to, an activity of a polypeptide of the present invention, including mature forms, as measured in a particular biological assay, with or without dose dependency. In the case where dose dependency does exist, it need not be identical to that of the polypeptide, but rather substantially similar to the dose-dependence in a given activity as compared to the polypeptide of the present
20 invention (i.e., the candidate polypeptide will exhibit greater activity or not more than about 25-fold less and, preferably, not more than about tenfold less activity, and most preferably, not more than about three-fold less activity relative to the polypeptide of the present invention.)

25 Polynucleotides and Polypeptides of the Invention

FEATURES OF PROTEIN ENCODED BY GENE NO: 1

30 The translation product of this gene shares sequence homology with a neurogenic secreted signaling protein, in addition to the human UDP-galactose:2-acetamido-2-deoxy-D-glucose3beta-galactosyltransferase (See Genbank Accession No. gnllPIDle1237254) which is thought to be vital in glycoprotein biosynthesis. In specific embodiments, polypeptides of the invention comprise the following amino acid
35 sequence: GLGPAQVALSLQGPA (SEQ ID NO:239), SSWMAGTQPRTSWWEMSS AKPCPTGTLRSNTSSHPQCTGPPTTHPMLVGEDMSCPEPQCGASRLSWKMNS

SPLMMSLWVCA (SEQ ID NO:240), QPRTSWWEMSSAKPCPTGTLRSN (SEQ ID NO:241), MSCPEPQCGASRLSWKMLNSSL (SEQ ID NO:242), WVALYIEG GMKYLTLVFLLGRAWRMTSPTRRSWAGSQPSRNSNTLGTWTKTSSSPFSMK WAWGQAATTQRCRCSSLSVRLKKSSVKSHWRMSSNSLLS (SEQ ID NO:243),
 5 GGMKYLTLVFLLGRAWRMTS (SEQ ID NO:244), SQPSRNSNTLGTWTKTS SSPFSMKW (SEQ ID NO:245), and/or TTQRCRCSSLSVRLKKSSVKSHWRMS (SEQ ID NO:246). Polynucleotides encoding these polypeptides are also encompassed by the invention. The gene encoding the disclosed cDNA is believed to reside on chromosome 12. Accordingly, polynucleotides related to this invention are useful as a
 10 marker in linkage analysis for chromosome 12.

This gene is expressed primarily in human fetal brain, epileptic frontal cortex and 12 week old early stage human.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a
 15 biological sample and for diagnosis of diseases and conditions which include, but are not limited to, neural or immune disorders, particularly neurodegenerative conditions such as epilepsy. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells,
 20 particularly of the central nervous system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., neural, immune, hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene
 25 expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:125 as residues: Ala-27 to Ser-38, Pro-43 to Asn-54, Thr-115 to Asp-121, Leu-225 to Val-232, Pro-247 to Gly-252, Arg-306 to Leu-311.

30 The tissue distribution in fetal brain tissue, combined with the homology to a neurogenic secreted signaling protein, in addition to the conserved UDP-galactose:2-acetamido-2-deoxy-D-glucose 3beta-galactosyltransferase protein indicates polynucleotides and polypeptides corresponding to this gene are useful for the treatment, detection, and/or prevention of a variety of neural disorders, particularly
 35 epilepsy and other disorders of the CNS. Moreover, polynucleotides and polypeptides corresponding to this gene are useful for the detection/treatment of neurodegenerative disease states, behavioral disorders, or inflammatory conditions which include, but are

not limited to Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, 5 paranoia, obsessive compulsive disorder, depression, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, elevated expression of this gene product in regions of the brain indicates that it plays a role in normal neural function. Potentially, this gene product is involved in synapse formation, 10 neurotransmission, learning, cognition, homeostasis, or neuronal differentiation or survival. In addition, the protein may show utility in the creation of novel therapeutics which depend upon the localizing benefits (cell and tissue specificity) of glycoproteins. This protein may also be used to produce physiologically active saccharide chains and variants, and for improvement of saccharide chains bound to physiologically active 15 proteins. Expression within fetal tissue and other cellular sources marked by proliferating cells indicates that this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis and treatment of cancer and other proliferative disorders. Similarly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation. Thus this protein may also be 20 involved in apoptosis or tissue differentiation and could again be useful in cancer therapy. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ 25 ID NO:11 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of 30 a-b, where a is any integer between 1 to 1257 of SEQ ID NO:11, b is an integer of 15 to 1271, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:11, and where b is greater than or equal to a + 14.

35 FEATURES OF PROTEIN ENCODED BY GENE NO: 2

In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: ASTLAQ

TTGTCKXXXSSRRARSRTQRXFQLRPDKRSAPSLQFIQAQEELSKENTGRQLA
AREAVLALEGSTQLTGPVTQVAASKTHCSGMALTASPVPVLGAAPAKXPTQ
5 NXPGQXGRAXXKVXTSWXXVATKVLHGLEVSTHLGKRKLSGRSWLPGP
ALHATPSQSHTQTGSQIVHPPQGEVREVGRGRGQPPAQPVHAHPSQQHPSPAH
LAGLSLWTGTA (SEQ ID NO:247), AMLETWRPGPSXGELATNSGQRASQDSQ
HSPPHVRAHLLISPLPAFPSMGGPAGRSAPXXLTETKSELQRLRRRQARASXS
XPAGEPGAGHSDSFNCVPTNGQPLRSCSLSKLRRSFLKRTQGDSWLPEKQSW
10 LWKAPPS (SEQ ID NO:248), SHQSHLINPASSAKGSAQLKAQPPAHVLGGT
GQEGPPPTADQPESPGWDPSSFTNGSSGPRALPTSVHPTLQQGAPCRRNWA
PCRGLVETRMLRRQLPHGTSKRDLGWASLQRGSPQETPQ (SEQ ID NO:249),
RPDKRSAPSLQFIQAQEELSKENTGRQLAAREAV (SEQ ID NO:250), ATPSQ
SHTQTGSQIVHPPQGEVREVGRGRGQPP (SEQ ID NO:251), QDSQHSPPHVR
15 AHLLISPLPAFPSMGGPA (SEQ ID NO:252), DSFNCVPTNGQPLRSCSL
KLRRSFLKR (SEQ ID NO:253), KGSWAQLKAQPPAHVLGGTGQEGPP (SEQ ID
NO 254:), KPSHQ (SEQ ID NO:256), and/or APSLLQFIQAQEELSKENTGRQLA
AR (SEQ ID NO:255). Polynucleotides encoding these polypeptides are also
encompassed by the invention.

20 This gene is expressed exclusively in adult human testis.

Therefore, polynucleotides and polypeptides of the invention are useful as
reagents for differential identification of the tissue(s) or cell type(s) present in a
biological sample and for diagnosis of diseases and conditions which include, but are
not limited to, reproductive disorders, particularly abnormalities of the testis, in addition
25 to impotence and infertility. Similarly, polypeptides and antibodies directed to these
polypeptides are useful in providing immunological probes for differential identification
of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells,
particularly of the male reproductive system, expression of this gene at significantly
higher or lower levels may be routinely detected in certain tissues or cell types (e.g.,
30 reproductive, testicular, adrogen regulated, and cancerous and wounded tissues) or
bodily fluids (e.g., lymph, serum, plasma, seminal fluid, urine, synovial fluid and
spinal fluid) or another tissue or cell sample taken from an individual having such a
disorder, relative to the standard gene expression level, i.e., the expression level in
healthy tissue or bodily fluid from an individual not having the disorder.

35 Preferred epitopes include those comprising a sequence shown in SEQ ID
NO:126 as residues: His-45 to Gly-56, Trp-62 to Tyr-68, His-94 to Trp-100.

The tissue distribution in testis indicates polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis, treatment, and/or prevention for abnormalities of the reproductive system. In addition, expression of this gene product in the testis may implicate this gene product in normal testicular function. This gene product may be useful in the treatment of male infertility, and/or could be used as a male contraceptive. Moreover, the protein product of this gene may be useful in the treatment, detection, and/or prevention of a variety of disorders related to androgen-regulated tissues, particularly the prostate gland. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:12 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1437 of SEQ ID NO:12, b is an integer of 15 to 1451, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:12, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 3

The translation product of this gene shares sequence homology with the human VAKTI precursor (See Genbank Accession No. gnllPIDle1311078 (AJ228139)), in addition to the ovoinhibitor and thrombin inhibitors, which are thought to be important in inhibition of protease activities. Contact of cells with supernatant expressing the product of this gene has been shown to increase the permeability of the plasma membrane of monocytes to calcium. Thus, it is likely that the product of this gene is involved in a signal transduction pathway that is initiated when the product binds a receptor on the surface of the plasma membrane of both immune cells, in addition to other cell-lines or tissue cell types. Thus, polynucleotides and polypeptides have uses which include, but are not limited to, activating monocytes. Binding of a ligand to a receptor is known to alter intracellular levels of small molecules, such as calcium, potassium and sodium, as well as alter pH and membrane potential. Alterations in small molecule concentration can be measured to identify supernatants which bind to

receptors of a particular cell. Moreover, when tested against NIH3T3 and U937 cell lines, supernatants removed from cells containing this gene activated the EGR1 (early growth response) and GAS (gamma activating sequence) promoter elements. Thus, it is likely that this gene activates fibroblasts or hematopoietic cells through the EGR1 and/or

5 JAK-STAT signal transduction pathway. EGR1 is a separate signal transduction pathway from Jak-STAT, genes containing the EGR1 promoter are induced in various tissues and cell types upon activation, leading the cells to undergo differentiation and proliferation. GAS is also a promoter element found upstream of many genes which are involved in the Jak-STAT pathway. The Jak-STAT pathway is a large, signal

10 transduction pathway involved in the differentiation and proliferation of cells. Therefore, activation of the Jak-STAT pathway, reflected by the binding of the GAS element, can be used to indicate proteins involved in the proliferation and differentiation of cells. In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: CSYRPQFPVDPRVRATCIVFN (SEQ ID NO:257), GTENLLA

15 PERTILSRAQMKGCMATPAPCVRSSSKQKKKKRKRKRVXQETKDNLRVQLPL XSCVVNXANPGKTDGFFAPERMTSPRAQMEKCMATPAPCVRPSFNKKKEQE QRLKEKLQRKSAVNFGTK (SEQ ID NO:258), LLAPERTILSRAQMKGCMAT PAPCVR (SEQ ID NO:259), PGKTDGFFAPERMTSPRAQMEKCM (SEQ ID NO:260), EQRLKEKLQRKSAVNFG (SEQ ID NO:261), KTLLENFSTQGTFFVAMH

20 PAVRATDWITLPCTKKPSISHLFFXFLAKILFSISSNSSFTLSLGIFSFXXQLST HCTLIAMRLPIRTKNRIIFPCASKSSISNKGPKSTAYILLWITALTFPFTFYTNL GPGFRILSTQCTSVVICFPICATNSFIIIRTDKIPISFSFFKIITIQLC WGSSLGSSC (SEQ ID NO:262), MHPAVRATDWITLPCTKKPSIS (SEQ ID NO:263), LIAMRLP IRTKNRIIFP (SEQ ID NO:264), SSISNKGPKSTAYILL WITALTFPFT (SEQ ID

25 NO:265), IIIRTDKIPISFSFFKIITIQLC (SEQ ID NO:266), NDGQCLAYNTTHY RERAMTSHARVSLGPSRDPLERPPFFFFFFFFFFFFKFEHTGTHGTLVSMHFAI WATDRIMLPGAYKCSIPHLVPKFTADFLCSFSFSLCSCSFLLKEGLTHGAGVA MHFSIWALDGVILSGAKKPSVFPGFAXFTTQLXKGSCSL RLSFVS (SEQ ID NO:267), CLAYNTTHYRERAMTSHARVSL (SEQ ID NO:268), GTLVSMHFAI

30 WATDRIMLPGAYKCSIPHLVP (SEQ ID NO:269), GVILSGAKK PSVFPGFAX FTTQLX (SEQ ID NO:270), KKASHMEQVLPICIFPSGPWMGSFSLXQKSRPF FLDLRXSLHNSXKEAVLLDCLLFLXXPSFFFFSSSAWKKKTSHMEQVLPCT FPSGPWIGLFSLVQASFPFLTSFRYSLQSSAYEVAFPDSSLFLARASAFFSSSFA WK (SEQ ID NO:271), CIFPSGPWMGSFSLXQKSRPFLLDLRXS (SEQ ID

35 NO:272), WIGLFSLVQASFPFLTSFRYSLQSSAYE (SEQ ID NO:273), NSAVN IKIRQRMEYFSVPEKMTLFFVQMGKCMATCVPCVKPTSKQKMKKRRLKHE LETKENLEKQPHMQSFAVNIESL (SEQ ID NO:274), IKIRQRMEYFSVPEKMTL

FVVQM (SEQ ID NO:275), and/or VKPTSKQKMKKRKRLKHELETKENL (SEQ ID NO:276). Polynucleotides encoding these polypeptides are also encompassed by the invention. The gene encoding the disclosed cDNA is believed to reside on chromosome 5. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 5.

This gene is expressed primarily in heart, tonsils, Hodgkin's lymphoma, neuroblastoma, leukocyte and lung.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, cardiovascular, immune, or hemodynamic disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the circulatory system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., cardiovascular, muscle, immune, hematopoietic, pulmonary, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, pulmonary surfactant or sputum, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:127 as residues: Ala-20 to Gln-27.

The tissue distribution in heart and immune cells and tissues, the homology to protease inhibitors, in addition to the detected calcium flux, EGR1, and GAS biological activities indicates polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of hemodynamic or vascular disorders, including hemorrhage, heart failure, and embolism, because proteases and their inhibitors are often involved in the cascades controlling hemodynamic controls. Protein may also show utility in the treatment, detection, and/or prevention of a variety of metabolic (i.e. cellular or physiological) and/or proliferative disorders in which aberrant regulation of a protease is thought to be involved, particularly in the premature activation of zymogens, for example. The secreted protein can also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions and as nutritional supplements. It may also have a very wide range of biological activities. Typical of these are cytokine, cell proliferation/differentiation modulating activity or induction of other cytokines;

immunostimulating/immunosuppressant activities (e.g. for treating human immunodeficiency virus infection, cancer, autoimmune diseases and allergy); regulation of hematopoiesis (e.g. for treating anemia or as adjunct to chemotherapy); stimulation or growth of bone, cartilage, tendons, ligaments and/or nerves (e.g. for treating wounds, stimulation of follicle stimulating hormone (for control of fertility); chemotactic and chemokinetic activities (e.g. for treating infections, tumors); hemostatic or thrombolytic activity (e.g. for treating hemophilia, cardiac infarction etc.); anti-inflammatory activity (e.g. for treating septic shock, Crohn's disease); as antimicrobials; for treating psoriasis or other hyperproliferative diseases; for regulation of metabolism, and behavior. Also contemplated is the use of the corresponding nucleic acid in gene therapy procedures. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:13 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome.

Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 2303 of SEQ ID NO:13, b is an integer of 15 to 2317, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:13, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 4

The translation product of this gene shares sequence homology with the ecotropic retrovirus receptor and the human cationic amino acid transporter-3 (See Genbank Accession No. gnllPIDle1198517) which are thought to be important in viral infections and amino acid and polyamine transport. In specific embodiments, polypeptides of the invention comprise the following amino acid sequence:

PRVRGTVVRLRQHRPSAYILVSTVLTLMPVWHS�DPDSALADAFYQRGYRWAG
FIVAAGSICA (SEQ ID NO:277), TVVRLRQHRPSAYILVSTVLTLMPV (SEQ ID
NO:278), WHSLDPDSALADAFYQRGYRWAGFIV (SEQ ID NO:279), TPSCSASS
SPCHALSMPWPPMGSSSRCLPMCTPGHRCLWRAPWRSGSSRPSWHCCWTWS

RWFSSCPLAHSWPTHSWPPVSLCCASRSLPRPAPQAQPALAP (SEQ ID NO:280), LSMPWPPMGSSSRCLPMCTPGHRC (SEQ ID NO:281), APWRSGSS RPSWHCCWTWSRWFSSCPL (SEQ ID NO:282), THSWPPVSLCCASRSL PRPAPQ (SEQ ID NO:283), AYILVSTVLTLMPVHSLDPDSALADAFYQRGYRW
 5 AGFIVAAGSICAMNTVLLSLLFSLP (SEQ ID NO:284), PWHSLDPDSALADAF YQRGYRWAGFIVAAGS (SEQ ID NO:285), RIVYAMAADGLFFQVFAHVHPRTQ VPV (SEQ ID NO:286), DLESLVQFLSLGTLLA (SEQ ID NO:287), YTFVATSII VLRFAQK (SEQ ID NO:288), LTKQQSSFSDDLQLVGTVHASVPEPGELKPA (SEQ ID NO:289), LRPYLGFLDGYSFGAVVTWALGVMLASAITIGCVLVFGNSTL
 10 HLPHWGYI (SEQ ID NO:290), PGAVVTWALGVMLASAITIGCVLVFGN (SEQ ID NO:291), GAHQQQYREDLFQIPMVPLIPALSIVLNICLMLKLSYLTWVRFISIW LLMGLAV (SEQ ID NO:292), MVPLIPALSIVLNICLMLKLSYLTWV (SEQ ID NO:293), and/or YFGYGIRHSENQRELPLNSTHYVVFPR (SEQ ID NO:294).
 Polynucleotides encoding these polypeptides are also encompassed by the invention.

15 This gene is expressed primarily in placenta and brain tissue.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, reproductive, neural, or metabolic disorders, in addition to viral
 20 infections. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the central nervous system and placenta, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g.,
 25 reproductive, neural, hepatic, metabolic, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, bile, amniotic fluid, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

30 Preferred epitopes include those comprising a sequence shown in SEQ ID NO:128 as residues: Gln-87 to Ser-99, Pro-102 to Phe-110, Gln-204 to Leu-211, Ser-262 to Glu-268, Pro-294 to His-305.

The tissue distribution in placenta, combined with the homology to a retroviral receptor and cationic amino acid transporters, indicates polynucleotides and
 35 polypeptides corresponding to this gene are useful for the diagnosis and intervention of viral infections, or diseases and malfunctions related to amino acid transport. Specifically, soluble forms of this protein or, polynucleotides of the present invention,

can be used to bind to retroviruses so as to prevent their entry and infection of susceptible cells. They can be used for therapy/prevention of HIV infection and certain forms of leukaemia. Polynucleotide or polypeptides of the present invention can be used to identify susceptibility to retroviral infection. Based upon the tissue distribution in the brain, polynucleotides and polypeptides corresponding to this gene are useful for the detection/treatment of neurodegenerative disease states, behavioral disorders, or inflammatory conditions which include, but are not limited to Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, depression, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, elevated expression of this gene product in regions of the brain indicates that it plays a role in normal neural function. Potentially, this gene product is involved in synapse formation, neurotransmission, learning, cognition, homeostasis, or neuronal differentiation or survival. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:14 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1458 of SEQ ID NO:14, b is an integer of 15 to 1472, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:14, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 5

In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: FPPSPAPPHSLPLRSWLWSRQMG (SEQ ID NO:295), GTSF RGMISTQPGSTPLASFKILALESADGHGGCSAGNDIGPYGERDDQQVFIQKVVP SASQLFVRLSSTGQRVCSVRSVDGSPTTAFTVLECEGSPAARLSAPALPAHWP

GQRQLGHVGPNNHRHGRPRPGPCRWPDGAR ADGTAGTL (SEQ ID NO:296),
 PGSTPLASFKILALESADGHGGCSAGNDI (SEQ ID NO:297), GERDDQQVF
 IQKVVPASQLFVRL (SEQ ID NO:298), RSVDGSPPTAFTVLECEGSPAARLS
 (SEQ ID NO:299), PALPAHWPGQRQLGHVGPNNHRHGRPR (SEQ ID NO:300),
 5 PFIPRRPWPEPGVPTGIREAPESPRTRASQGIMAAALFKKEVSLSFILGGVRG
 VPRPLEGHGAGVGGRRRSGPLRTSSWQRSTKLPPRRRRASACGGLGLPRWP
 DKEVLLEAEWRLVREMRGEGLGRQPHEGAERSRRGQLTVFQLFHQLLLRQATC
 RGLA EAVHGGG (SEQ ID NO:301), PGVPTGIREAPESPRTRASQGIMAAALF
 KKEV (SEQ ID NO:302), FILGGVRGVPRPLEGHGAGVGGRRRSGP (SEQ ID
 10 NO:303), GLPRWPDKEVLLEAEWRLVREMRG (SEQ ID NO:304), and/or GAER
 SRRGQLTVFQLFHQLLLRQ (SEQ ID NO:305). Polynucleotides encoding these
 polypeptides are also encompassed by the invention.

This gene is expressed primarily in human fetal kidney, and to a lesser extent, in
 thymus and bone marrow cell line (RS4;11).

15 Therefore, polynucleotides and polypeptides of the invention are useful as
 reagents for differential identification of the tissue(s) or cell type(s) present in a
 biological sample and for diagnosis of diseases and conditions which include, but are
 not limited to, developmental, metabolic, immune or hematopoietic disorders.
 Similarly, polypeptides and antibodies directed to these polypeptides are useful in
 20 providing immunological probes for differential identification of the tissue(s) or cell
 type(s). For a number of disorders of the above tissues or cells, particularly of the
 immune system, expression of this gene at significantly higher or lower levels may be
 routinely detected in certain tissues or cell types (e.g., developmental, metabolic,
 immune, hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g.,
 25 lymph, serum, amniotic fluid, plasma, urine, synovial fluid and spinal fluid) or another
 tissue or cell sample taken from an individual having such a disorder, relative to the
 standard gene expression level, i.e., the expression level in healthy tissue or bodily
 fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID
 30 NO:129 as residues: Thr-17 to Trp-26, Pro-54 to Trp-61, Ala-65 to Arg-74, Pro-142 to
 Leu-147, Pro-158 to Ala-165.

The tissue distribution in thymus and bone marrow cell lines indicates
 polynucleotides and polypeptides corresponding to this gene are useful for the
 diagnosis, treatment, and/or prevention of immune disorders involving stem cells.
 35 Moreover, polynucleotides and polypeptides corresponding to this gene are useful for
 the treatment and diagnosis of hematopoietic related disorders such as anemia,
 pancytopenia, leukopenia, thrombocytopenia or leukemia since stromal cells are

important in the production of cells of hematopoietic lineages. The uses include bone marrow cell ex-vivo culture, bone marrow transplantation, bone marrow reconstitution, radiotherapy or chemotherapy of neoplasia. The gene product may also be involved in lymphopoiesis, therefore, it can be used in immune disorders such as infection, inflammation, allergy, immunodeficiency etc. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Alternatively, the expression within fetal kidney indicates that this gene or gene product could be used in the treatment and/or detection of kidney diseases including renal failure, nephritis, renal tubular acidosis, proteinuria, pyuria, edema, pyelonephritis, hydronephritis, nephrotic syndrome, crush syndrome, glomerulonephritis, hematuria, renal colic and kidney stones, in addition to Wilm's Tumor Disease, and congenital kidney abnormalities such as horseshoe kidney, polycystic kidney, and Falconi's syndrome. Moreover, the expression within fetal tissue indicates that this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis and treatment of cancer and other proliferative disorders. Similarly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation. Thus this protein may also be involved in apoptosis or tissue differentiation and could again be useful in cancer therapy. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:15 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1002 of SEQ ID NO:15, b is an integer of 15 to 1016, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:15, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 6

In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: HASAHASAHASGCGA (SEQ ID NO:306), QGVGVADGG

LERQRVDAGARLGHMGQPVAFSTRQLHLALPAPGTAGVTVPHPHAREGVV
 GDLPLVPDAEDPTVGVPAGEGLLVLGHVVERAELILVRGLHQAEALARESEEMH
 GSRHG (SEQ ID NO:307), EGGLERQRVDAGARLGHMGQPVAFS (SEQ ID
 NO:308), LALPAPGTAGVTVPHPHAREGVVGDPLV (SEQ ID NO:309), PAEG
 5 LLVLGHVVERAELILVRGLHQAEA (SEQ ID NO:310), HLFKFFYTIAFMQWF
 TEFMELFLSVWELIKTXNLFCVCFSEHKPGQLVPAGPTSQLLCRALGRVH
 LCSPTTRSQTPTQSWVTPQLLWRLGSGRLVAQVLQVGSFCGPRVGDAVLGEQT
 FQPFDLL (SEQ ID NO:311), AFMQWFTEFMELFLSVWELIKTXNLFCVC (SEQ
 ID NO:312), and/or RSQTPTQSWVTPQLLWRLGSGRLVAQ (SEQ ID NO:313).

- 10 Polynucleotides encoding these polypeptides are also encompassed by the invention.
 The gene encoding the disclosed cDNA is believed to reside on chromosome 16.
 Accordingly, polynucleotides related to this invention are useful as a marker in linkage
 analysis for chromosome 16.

- This gene is expressed primarily in human infant brain, and to a lesser extent, in
 15 adult brain and lung.

- Therefore, polynucleotides and polypeptides of the invention are useful as
 reagents for differential identification of the tissue(s) or cell type(s) present in a
 biological sample and for diagnosis of diseases and conditions which include, but are
 not limited to, neural, developmental, or pulmonary disorders. Similarly, polypeptides
 20 and antibodies directed to these polypeptides are useful in providing immunological
 probes for differential identification of the tissue(s) or cell type(s). For a number of
 disorders of the above tissues or cells, particularly of the central nervous system
 (CNS), expression of this gene at significantly higher or lower levels may be routinely
 detected in certain tissues or cell types (e.g., neural, developmental, pulmonary, and
 25 cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine,
 amniotic fluid, synovial fluid and spinal fluid) or another tissue or cell sample taken
 from an individual having such a disorder, relative to the standard gene expression
 level, i.e., the expression level in healthy tissue or bodily fluid from an individual not
 having the disorder.

- 30 Preferred epitopes include those comprising a sequence shown in SEQ ID
 NO:130 as residues: Ser-47 to Pro-57, Ser-77 to Glu-82, Thr-90 to Trp-98, Arg-124 to
 Lys-137, Ala-183 to Glu-192, Lys-220 to Gln-229, Asn-244 to Arg-258, Thr-271 to
 Asn-278, Glu-285 to Gly-297.

- The tissue distribution in infant and adult brain indicates polynucleotides and
 35 polypeptides corresponding to this gene are useful for the diagnosis, treatment, and/or
 prevention of diseases of the CNS, such as mental retardation, schizophrenia,
 Alzheimer's disease, paranoia, depression, and mania. Moreover, polynucleotides and

polypeptides corresponding to this gene are useful for the detection/treatment of neurodegenerative disease states, behavioral disorders, or inflammatory conditions which include, but are not limited to Parkinson's Disease, Huntington's Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, dementia, obsessive compulsive disorder, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, elevated expression of this gene product in regions of the brain indicates that it plays a role in normal neural function. Potentially, this gene product is involved in synapse formation, neurotransmission, learning, cognition, homeostasis, or neuronal differentiation or survival. The protein product of this gene may also show utility in the detection, treatment, and/or prevention of a variety of pulmonary disorders, particularly those related to disorders of the mucosal or endothelial tissues. The expression within fetal tissue indicates that this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis and treatment of cancer and other proliferative disorders. Similarly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation. Thus this protein may also be involved in apoptosis or tissue differentiation and could again be useful in cancer therapy. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:16 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1225 of SEQ ID NO:16, b is an integer of 15 to 1239, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:16, and where b is greater than or equal to a + 14.

35

FEATURES OF PROTEIN ENCODED BY GENE NO: 7

The gene encoding the disclosed cDNA is believed to reside on chromosome 17.

Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 17. In specific embodiments, polypeptides of the invention

5 comprise the following amino acid sequence:

GAWGVEVVAVGSKAGCLVYQLCDLKQITFFFRASVCLSV (SEQ ID NO:314),
 PASLGSSWGQKLRRGGTRKSFQELSPSSAPPACLPQPPASTWLSSWPRPPCW
 PPMCSWALGXCFPATGQWLPTSCCLWWCPDAGGRQKHFRSRWXTSWETW
 QPYLTGLISSVLRAXRPDSYLQFRSLXQXXLCCAFVIALGGGCFLLTALYLER
 10 DETRAWQXVTGTPDSNDVDSNDLERQGLLSGXGASTEEL (SEQ ID NO:315), L
 RGGTRKSFQELSPSSAPPACLPQPP (SEQ ID NO:316), ATGQWLPTSCCLW
 WCPDAGGRQKHFRSR (SEQ ID NO:317), GGCFLLTALYLERDETRAWQXV
 (SEQ ID NO:318), APHLRLQPACHSPLPLPGSRPGPDHPAGLLCVPGPWG
 ASVLQLGSGCRHPAVCGGAQMPGDGRSTSDHGGXHPGXPGSPISQDLVSLVSC
 15 GPXALTPICSASAAXXXXXCAAPLSPWGAAASC (SEQ ID NO:319),
 PACHSPLPLPGSRPGPDHPAGLLCV (SEQ ID NO:320), and/or
 SGRHPAVCGGAQMPGDGRSTSDHGG (SEQ ID NO:321). Polynucleotides

encoding these polypeptides are also encompassed by the invention.

This gene is expressed primarily in pineal gland and thymus.

20 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune, endocrine, emotional or behavior disorders, in addition to autoimmune diseases. Similarly, polypeptides and antibodies directed to these
 25 polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., immune, endocrine, hematopoietic, neural, and cancerous and wounded tissues) or bodily fluids
 30 (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID
 35 NO:131 as residues: Asp-18 to Gln-27, Arg-44 to Asn-49.

The tissue distribution in thymus indicates polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis, treatment, and/or prevention of

immune and autoimmune diseases, such as lupus, neutropenia, transplant rejection, and inflammatory diseases. Moreover, the expression within pineal gland indicates the protein product of this gene may be useful in disorders associated with biological clock aberrations, emotional distress, lethargy, or metabolic conditions. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:17 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1391 of SEQ ID NO:17, b is an integer of 15 to 1405, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:17, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 8

20

When tested against K562 cell lines, supernatants removed from cells containing this gene activated the ISRE (interferon-sensitive responsive element) promoter element. Thus, it is likely that this gene activates leukemia cells, or more generally, immune or hematopoietic cells, through the JAK-STAT signal transduction pathway. ISRE is a promoter element found upstream in many genes which are involved in the Jak-STAT pathway. The Jak-STAT pathway is a large, signal transduction pathway involved in the differentiation and proliferation of cells. Therefore, activation of the Jak-STAT pathway, reflected by the binding of the ISRE element, can be used to indicate proteins involved in the proliferation and differentiation of cells. In specific embodiments, polypeptides of the invention comprise the following amino acid sequence:

GLKVMEICSLTFLEATNLQSRCQQAMLPLKALRKNPFLLLPSFDGCCQSLA
FPGLWLQHSNLCLNHHMTFLVYLLCVSVFKYFFPFSCCTYTSHWI (SEQ ID
NO:322), ICSLTFLEATNLQSRCQQAMLPL (SEQ ID NO:323), and/or GLWLQHS
NLCLNHHMTFLVYLLCVSV (SEQ ID NO:324). Polynucleotides encoding these
polypeptides are also encompassed by the invention.

This gene is expressed primarily in IL-1 and LPS induced neutrophils.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune or hematopoietic disorders, particularly inflammation. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., immune, hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:132 as residues: Ser-45 to His-50, His-52 to Ile-57, Lys-67 to His-81.

The tissue distribution in neutrophils, combined with the detected ISRE biological activity indicates polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis, treatment, and/or prevention of inflammatory disorders, such as psoriasis, inflammatory bowel disease, rheumatoid arthritis, and sepsis. Moreover, polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of hematopoietic related disorders such as anemia, pancytopenia, leukopenia, thrombocytopenia or leukemia since stromal cells are important in the production of cells of hematopoietic lineages. The uses include bone marrow cell ex-vivo culture, bone marrow transplantation, bone marrow reconstitution, radiotherapy or chemotherapy of neoplasia. The gene product may also be involved in lymphopoiesis, therefore, it can be used in immune disorders such as infection, inflammation, allergy, immunodeficiency etc. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:18 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome.

Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1520 of SEQ ID NO:18, b is an integer of 15 to 1534, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:18, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 9

10 In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: PFPLLPPKRRGLLYHLLIQKSTLGLVWVFREHLDSRSQ (SEQ ID NO:325), RGXPSWPMHTLVYAQHSTTHTPLIQPQWTQVIDQPPGITHQFCVR XCXCPTLESCVQECVTRSXRKPTTGVPGPQRLA (SEQ ID NO:326), TPLIQPQW TQVIDQPPGITHQFCV (SEQ ID NO:327), ALGPSQTCDLVWLVAKPSFFRGPQ
15 GIHYFSLWRRKPLSHWVSIWQLQGQETMPAMLRSRPAGQATVATGPPRGSPS PQDLPSYHRKQVESSHRHSWEPASQSQ (SEQ ID NO:328), CDLDVWLVAKPSF FRGPQGIHYFSLWRR (SEQ ID NO:329), and/or AGQATVATGPPRGSPSPQDLP SYHRKQV (SEQ ID NO:330). Polynucleotides encoding these polypeptides are also encompassed by the invention.

20 This gene is expressed primarily in synovial fibroblasts, and to a lesser extent, in endothelial cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, skeletal or vascular disorders, particularly arthritis. Similarly,
25 polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the skeletal system, expression of this gene at significantly higher or lower levels may be routinely detected
30 in certain tissues or cell types (e.g., skeletal, vascular, endothelial, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

35 The tissue distribution in synovial fibroblasts indicates polynucleotides and polypeptides corresponding to this gene are useful for treatment of arthritis. Moreover, polynucleotides and polypeptides corresponding to this gene are useful in the detection

and treatment of disorders and conditions afflicting the skeletal system, in particular osteoporosis, bone cancer, as well as, disorders afflicting connective tissues (e.g. arthritis, trauma, tendonitis, chondromalacia and inflammation), such as in the diagnosis or treatment of various autoimmune disorders such as rheumatoid arthritis, lupus, scleroderma, and dermatomyositis as well as dwarfism, spinal deformation, and specific joint abnormalities as well as chondrodysplasias (i.e. spondyloepiphyseal dysplasia congenita, familial osteoarthritis, Atelosteogenesis type II, metaphyseal chondrodysplasia type Schmid). The protein product of this gene may also be useful for the detection, treatment, and/or prevention of a variety of vascular disorders which include, but are not limited to, atherosclerosis, embolism, stroke, microvascular disease, or aneurysm. The protein may also be useful in the treatment of integumentary disorders, particularly those related to aberrations in the extracellular matrix or lamina. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:19 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1219 of SEQ ID NO:19, b is an integer of 15 to 1233, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:19, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 10

In specific embodiments, polypeptides of the invention comprise the following amino acid sequence:

XGDTXTQNSRHDTPLIDYYRESCTLQYRPEFPGRPTRPRGSCPQYPGPAIPRT
 SWALGEGDAAPRGAHH PRRADVPLG (SEQ ID NO:331), YRESCTLQYRPEFPG
 RPTRPRGSCPQYPGP (SEQ ID NO:332), GKLYAAVPSGIPGSTHASARLMPPVS
 RSSYSEDIVGSRRRRRSSSGSPSPQSRCSSWDGCSRSHSRGREGXRPPWSEL
 DVGALYPFSRSGSRGRLPRFRNYAFASSWSTS YSGYRYHRALLCRRTAVSGR
 LREGREPSAEEAEGEREDWGIGSA (SEQ ID NO:333), SGIPGSTHASARLMPP

VSRSSYS (SEQ ID NO:334), GCSRSHSRGREGXRPPWSELV GALYPFS (SEQ ID NO:335), TAVSGRLREGREPSAEEAEGEREDW (SEQ ID NO:336), RIRKAA VQIPTRKNIGXRRPVVQETRKKERISRLKESIGNILIVTVTQSLTQILTLMMI KRELKPRRKRRKRNTKQXKRRIRKPKKNPVTQAVKTQKRTCQKLPGMEQ
 5 PNVADTMDLIGPEAPINTYLFKMKNL (SEQ ID NO:337), TRKKERISRLKESIGNILIVTVTQSLTQ (SEQ ID NO:338), and/or VKTQKRTCQKLPGMEQPNVA DTMDLIGP (SEQ ID NO:339). Polynucleotides encoding these polypeptides are also encompassed by the invention. The gene encoding the disclosed cDNA is believed to reside on chromosome 6. Accordingly, polynucleotides related to this invention are
 10 useful as a marker in linkage analysis for chromosome 6.

This gene is expressed primarily in retina.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are
 15 not limited to, visual disorders, particularly retinopathy. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the ocular system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell
 20 types (e.g., cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, vitreous humor, aqueous humor, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

25 The tissue distribution in retina indicates polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis, treatment, and/or prevention of diseases of the retina, for example, diabetic retinopathy. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

30 Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:20 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome.

35 Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1076 of SEQ ID NO:20, b is an integer of 15

to 1090, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:20, and where b is greater than or equal to a + 14.

5

FEATURES OF PROTEIN ENCODED BY GENE NO: 11

When tested against NIH3T3 cell lines, supernatants removed from cells containing this gene activated the EGR1 (early growth response gene 1) promoter element. Thus, it is likely that this gene activates fibroblast cells, or more generally, other cells of the integumentary system, through the EGR1 signal transduction pathway. EGR1 is a separate signal transduction pathway from Jak-STAT. Genes containing the EGR1 promoter are induced in various tissues and cell types upon activation, leading the cells to undergo differentiation and proliferation. In specific embodiments, polypeptides of the invention comprise the following amino acid sequence:

LPFTLKPKMVKIPFSSRLINNNLQYIDCILSLKRCEEILLMWHGLLLCLASVFLE
 LRGDRPPLLASLLEPHKMPLHSSSL (SEQ ID NO:340), LKPKMVKIPFSSRLIN
 NNLQYIDC (SEQ ID NO:341), SLKRCEEILLMWHGLLLCLASVF (SEQ ID
 NO:342), LRGDRPPLLASLLEPHKMPLH (SEQ ID NO:343), LQMHTGSGFKGK
 SCEVAFYVAQAEPGEGAYLHGAQETQKQIEADHATLRGSPHSVSKTKYNLY
 IANYLLAWRKMES (SEQ ID NO:344), CEVAFYVAQAEPGEGAYLH (SEQ ID
 NO:345), and/or ATLRGSPHSVSKTKYNLYIANY (SEQ ID NO:346).
 Polynucleotides encoding these polypeptides are also encompassed by the invention.

25 This gene is expressed primarily in human ovarian cancer.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, reproductive disorders, particularly cancers, such as ovarian cancer.

30 Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the female reproductive system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., reproductive, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual

35 having such a disorder, relative to the standard gene expression level, i.e., the

expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in human ovarian cancer tissue indicates polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis, treatment, and/or prevention of ovarian cancer. Moreover, the expression within cellular sources marked by proliferating cells indicates that this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis and treatment of cancer and other proliferative disorders. Similarly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation. Thus this protein may also be involved in apoptosis or tissue differentiation and could again be useful in cancer therapy. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:21 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 668 of SEQ ID NO:21, b is an integer of 15 to 682, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:21, and where b is greater than or equal to a + 14.

25 FEATURES OF PROTEIN ENCODED BY GENE NO: 12

Contact of cells with supernatant expressing the product of this gene increases the permeability of the plasma membrane of myeloid leukemia cells to calcium. Thus, it is likely that the product of this gene is involved in a signal transduction pathway that is initiated when the product binds a receptor on the surface of immune or hematopoietic cells, in addition to other cells or cell types. Thus, polynucleotides and polypeptides have uses which include, but are not limited to, activating myeloid cells. Binding of a ligand to a receptor is known to alter intracellular levels of small molecules, such as calcium, potassium and sodium, as well as alter pH and membrane potential. Alterations in small molecule concentration can be measured to identify supernatants which bind to receptors of a particular cell. In specific embodiments, polypeptides of the invention comprise the following amino acid sequence:

LSASLLDRYPASESNNYIFNFVLYMLHFLAGTLFSLFPDFELSPRSATLFPDLR
 TVQLSSRPHL (SEQ ID NO:347), LLDRYPASESNNYIFNFVLYMLH (SEQ ID
 NO:348), FPDFELSPRSATLFPDLRTV (SEQ ID NO:349), NGGFYDVSKQAG
 LIEFLCIIFYPMAHVICGSRFTIVRTIPVHYVGEYFIKSSIWILYRINERTATKK
 5 AASDFQKNFRCFLDAF (SEQ ID NO:350), KQAGLIEFLCIIFYPMAH (SEQ ID
 NO:351), and/or YFIKSSIWILYRINERTATKAA (SEQ ID NO:352).

Polynucleotides encoding these polypeptides are also encompassed by the invention.

This gene is expressed primarily in anergic T-cells.

Therefore, polynucleotides and polypeptides of the invention are useful as
 10 reagents for differential identification of the tissue(s) or cell type(s) present in a
 biological sample and for diagnosis of diseases and conditions which include, but are
 not limited to, immune or hematopoietic disorders, particularly immunodeficiencies and
 inflammatory disorders. Similarly, polypeptides and antibodies directed to these
 polypeptides are useful in providing immunological probes for differential identification
 15 of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells,
 particularly of the immune system, expression of this gene at significantly higher or
 lower levels may be routinely detected in certain tissues or cell types (e.g., immune,
 hematopoietic cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum,
 plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken
 20 from an individual having such a disorder, relative to the standard gene expression
 level, i.e., the expression level in healthy tissue or bodily fluid from an individual not
 having the disorder.

The tissue distribution in anergic T-cells, combined with the detected calcium
 flux biological activity, indicates polynucleotides and polypeptides corresponding to
 25 this gene are useful for the diagnosis and treatment of immune disorders, particularly
 those involving anergic T-cells. Moreover, the secreted protein can also be used to
 determine biological activity, to raise antibodies, as tissue markers, to isolate cognate
 ligands or receptors, to identify agents that modulate their interactions, and as
 nutritional supplements. It may also have a very wide range of biological activities.
 30 Typical of these are cytokine, cell proliferation/differentiation modulating activity or
 induction of other cytokines; immunostimulating/immunosuppressant activities (e.g. for
 treating human immunodeficiency virus infection, cancer, autoimmune diseases and
 allergy); regulation of hematopoiesis (e.g. for treating anaemia or as adjunct to
 chemotherapy); stimulation or growth of bone, cartilage, tendons, ligaments and/or
 35 nerves (e.g. for treating wounds, stimulation of follicle stimulating hormone (for
 control of fertility); chemotactic and chemokinetic activities (e.g. for treating infections,
 tumors); hemostatic or thrombolytic activity (e.g. for treating haemophilia, cardiac

infarction etc.); anti-inflammatory activity (e.g. for treating septic shock, Crohn's disease); as antimicrobials; for treating psoriasis or other hyperproliferative diseases; for regulation of metabolism, and behaviour. Also contemplated is the use of the corresponding nucleic acid in gene therapy procedures. Protein, as well as, antibodies
5 directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:22 and may have been publicly available prior to conception of the present
10 invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 756 of SEQ ID NO:22, b is an integer of 15 to
15 770, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:22, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 13

20

In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: SPRXGRXFXTSRKQISGFLEFD (SEQ ID NO:353). Polynucleotides encoding these polypeptides are also encompassed by the invention.

This gene is expressed primarily in bone marrow.

25

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune or hematopoietic disorders, particularly leukemia or multiple myeloma. Similarly, polypeptides and antibodies directed to these polypeptides are
30 useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., immune, hematopoietic cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine,
35 synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the

expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in bone marrow tissue indicates polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis, treatment, and/or prevention of some types of leukemia, and other disorders involving bone marrow tissues or cells. Moreover, polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of hematopoietic related disorders such as anemia, pancytopenia, leukopenia, thrombocytopenia or leukemia since stromal cells are important in the production of cells of hematopoietic lineages. The uses include bone marrow cell ex-vivo culture, bone marrow transplantation, bone marrow reconstitution, radiotherapy or chemotherapy of neoplasia. The gene product may also be involved in lymphopoiesis, therefore, it can be used in immune disorders such as infection, inflammation, allergy, immunodeficiency etc. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:23 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 551 of SEQ ID NO:23, b is an integer of 15 to 565, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:23, and where b is greater than or equal to a + 14.

30 FEATURES OF PROTEIN ENCODED BY GENE NO: 14

In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: MKHAAFGLIPLVKEIYRYLKI KSKLLIGSGKCQLQPEWL QTSLINSSLLMDWLTPY (SEQ ID NO:354), IYRYLKI KSKLLIGSGKCQLQPE WLQTS (SEQ ID NO:355), QLGLPWDQSKGPRKNGLSMCGSVYSTIWSLIA SRREETIRVIVLYIQSPNINTRHISKRGLN KALTNP (SEQ ID NO:356), SKGPR

KNGLSMCGSVYSTIWS (SEQ ID NO:357), and/or QSPNINTRHISKRGLNK (SEQ ID NO:358). Polynucleotides encoding these polypeptides are also encompassed by the invention.

5 This gene is expressed primarily in adult retina, and to a lesser extent, in 12 week old early stage human.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, visual or developmental disorders, particularly retinopathy. Similarly,
10 polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the ocular system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., visual, developmental cancerous and wounded
15 tissues) or bodily fluids (e.g., lymph, serum, amniotic fluid, aqueous humor, vitreous humor, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

20 The tissue distribution in human retina indicates polynucleotides and polypeptides corresponding to this gene are useful for treatment of retinopathy, and other disorders involving the visual system. Moreover, the expression within embryonic tissue indicates that this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis and treatment of cancer and other
25 proliferative disorders. Similarly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation. Thus this protein may also be involved in apoptosis or tissue differentiation and could again be useful in cancer therapy. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

30 Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:24 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome.

35 Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1342 of SEQ ID NO:24, b is an integer of 15

to 1356, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:24, and where b is greater than or equal to a + 14.

5 FEATURES OF PROTEIN ENCODED BY GENE NO: 15

The translation product of this gene has homology to the conserved human non-differentiated blood cell tyrosine kinase receptor fragment (See Genbank Accession No. R76466) which is thought to be important in signalling essential cellular pathways. In
 10 specific embodiments, polypeptides of the invention comprise the following amino acid sequence: HPQTSAGGFPLHQGLPTVS (SEQ ID NO:359), PSWFPELSPWPLKTL KKRRQMRLRRRGRLCRLSPATTTTADTCRCPARSYRGSGRRPACAQDSPAPPS RPTTRRAWEKCALRPKRAAQWSTGVPPSPRSSTTGCCFGTAAXCAEGARR (SEQ ID NO:360), TTTADTCRCPARSYRGSGRRPA (SEQ ID NO:361), and/or
 15 PSRPTTRRAWEKCALRPKRAAQWST (SEQ ID NO:362). Polynucleotides encoding these polypeptides are also encompassed by the invention.

This gene is expressed primarily in human fetal epithelium.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a
 20 biological sample and for diagnosis of diseases and conditions which include, but are not limited to, developmental, integumentary, immune, or hematopoietic disorders, particularly skin cancer. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells,
 25 particularly of the integumental system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., developmental, integumentary, immune, hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, amniotic fluid, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having
 30 such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:139 as residues: Gln-26 to Ala-39, Cys-48 to His-55.

The tissue distribution in human fetal epithelium, combined with the homology
 35 to a conserved tyrosine kinase receptor, indicates polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis, treatment, and/or prevention of skin cancer, or other disorders related to the integument, particularly proliferative

conditions. Similarly, polynucleotides and polypeptides corresponding to this gene are useful for the treatment, diagnosis, and/or prevention of various skin disorders including congenital disorders (i.e. nevi, moles, freckles, Mongolian spots, hemangiomas, port-wine syndrome), integumentary tumors (i.e. keratoses, Bowen's disease, basal cell carcinoma, squamous cell carcinoma, malignant melanoma, Paget's disease, mycosis fungoides, and Kaposi's sarcoma), injuries and inflammation of the skin (i.e. wounds, rashes, prickly heat disorder, psoriasis, dermatitis), atherosclerosis, urticaria, eczema, photosensitivity, autoimmune disorders (i.e. lupus erythematosus, vitiligo, dermatomyositis, morphea, scleroderma, pemphigoid, and pemphigus), keloids, striae, erythema, petechiae, purpura, and xanthelasma. In addition, such disorders may predispose increased susceptibility to viral and bacterial infections of the skin (i.e. cold sores, warts, chickenpox, molluscum contagiosum, herpes zoster, boils, cellulitis, erysipelas, impetigo, tinea, athlete's foot, and ringworm). Moreover, the protein product of this gene may also be useful for the treatment or diagnosis of various connective tissue disorders such as arthritis, trauma, tendonitis, chondromalacia and inflammation, autoimmune disorders such as rheumatoid arthritis, lupus, scleroderma, and dermatomyositis as well as dwarfism, spinal deformation, and specific joint abnormalities as well as chondrodysplasias (ie. spondyloepiphyseal dysplasia congenita, familial osteoarthritis, Atelosteogenesis type II, metaphyseal chondrodysplasia type Schmid). Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:25 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 603 of SEQ ID NO:25, b is an integer of 15 to 617, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:25, and where b is greater than or equal to a + 14.

35 FEATURES OF PROTEIN ENCODED BY GENE NO: 16

When tested against PC12 cell lines, supernatants removed from cells containing this gene activated the EGR1 (early growth response gene 1) pathway. Thus, it is likely that this gene activates sensory neuron cells, or generally other cells or cell types, particularly immune cells, through the EGR1 signal transduction pathway. EGR1 is a separate signal transduction pathway from Jak-STAT, genes containing the EGR1 promoter are induced in various tissues and cell types upon activation, leading the cells to undergo differentiation and proliferation. In specific embodiments, polypeptides of the invention comprise the following amino acid sequence:

ARGVLNLRNRFECFSIIETV (SEQ ID NO:363), IGQLVMKSICHFQRLLSVAI
DFASQFLKNYIFSSTHSSKAGFSVVCSLPKWLYTDGMEMVLKITHKLSF (SEQ
ID NO:364), and/or QRLLSVAIDFASQFLKNYIFSSTH (SEQ ID NO:365).

Polynucleotides encoding these polypeptides are also encompassed by the invention. The gene encoding the disclosed cDNA is believed to reside on chromosome 8. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 8.

This gene is expressed primarily in fetal liver, and to a lesser extent, in resting T-cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune, hematopoietic, or hepatic disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., immune, hematopoietic, hepatic, metabolic, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, amniotic fluid, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in fetal liver and resting T-cells, combined with the detected EGR1 biological activity indicates polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis, treatment, and/or prevention of immune disorders involving T-cells, and more generally, hematopoietic conditions. Moreover, polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of hematopoietic related disorders such as anemia.

pancytopenia, leukopenia, thrombocytopenia or leukemia since stromal cells are important in the production of cells of hematopoietic lineages. The uses include bone marrow cell ex-vivo culture, bone marrow transplantation, bone marrow reconstitution, radiotherapy or chemotherapy of neoplasia. The gene product may also be involved in lymphopoiesis, therefore, it can be used in immune disorders such as infection, inflammation, allergy, immunodeficiency etc. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Additionally, expression within fetal tissue indicates that this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis and treatment of cancer and other proliferative disorders. Similarly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation. Thus this protein may also be involved in apoptosis or tissue differentiation and could again be useful in cancer therapy. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:26 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 634 of SEQ ID NO:26, b is an integer of 15 to 648, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:26, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 17

30

In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: LMKTASRMLLLE (SEQ ID NO:366). Polynucleotides encoding these polypeptides are also encompassed by the invention.

This gene is expressed primarily in CD34-positive T cells from cord blood, and to a lesser extent, in anergic T cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a

biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune or hematopoietic disorders, particularly in immune system maturation and hematopoietic development. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for

5 differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., immune, hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, amniotic fluid, urine, synovial fluid and spinal

10 fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:141 as residues: Ile-46 to Tyr-56.

15 The tissue distribution in CD34-positive T cells and anergic T cells. indicates polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis, treatment, and/or prevention of diseases involving hematopoietic development and stem cell maturation, including protection of stem cells from chemotherapy, immunosuppression during transplant rejection, and neutropenia.

20 Moreover, this gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions. Therefore it may be also used as an agent for immunological disorders

25 including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune

30 infertility, lens tissue injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's disease, scleroderma and tissues. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies

35 directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:27 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1374 of SEQ ID NO:27, b is an integer of 15 to 1388, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:27, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 18

When tested against U937 and Jurket cell lines, supernatants removed from cells containing this gene activated the GAS (gamma activating sequence) pathway. Thus, it is likely that this gene activates myeloid and T-cells, or more generally cells of immune or hematopoietic origin, through the JAK-STAT signal transduction pathway. GAS is a promoter element found upstream of many genes which are involved in the Jak-STAT pathway. The Jak-STAT pathway is a large, signal transduction pathway involved in the differentiation and proliferation of cells. Therefore, activation of the Jak-STAT pathway, reflected by the binding of the GAS element, can be used to indicate proteins involved in the proliferation and differentiation of cells. In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: ATXWDXPGCRNSARGERLHVGDAPW (SEQ ID NO:367), ARDER REVLKTLMLRLSTQRPQAFLPSQSWFVRLQKAGEGALKQENSLTIQNCLLCL PRVHRQRPTPPQPQRGNTASVLQTSTEHLPRAAVLLVPNSCSPGXPTXLLSS (SEQ ID NO:368), ERREVLKTLMLRLSTQRPQAFLP (SEQ ID NO:369), GALKQEN SLTIQNCLLCLPRVHRQR (SEQ ID NO:370), and/or SVLQTSTEHLPRAAVLLVP NS (SEQ ID NO:371). Polynucleotides encoding these polypeptides are also encompassed by the invention.

This gene is expressed primarily in activated human neutrophils.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune or hematopoietic disorders, such as neutropenia. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing

immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., immune, hematopoietic cancerous and wounded
5 tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID
10 NO:142 as residues: Val-25 to Gly-33.

The tissue distribution in activated neutrophils, combined with the detected GAS biological activity, indicates polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis, treatment, and/or prevention of immune disorders involving neutrophils. Moreover, this gene product may be involved in the regulation of
15 cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS,
20 leukemia, rheumatoid arthritis, granulomatous disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus
25 erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's disease, scleroderma and tissues. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or
30 immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:28 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the
35 scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of

a-b, where a is any integer between 1 to 602 of SEQ ID NO:28, b is an integer of 15 to 616, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:28, and where b is greater than or equal to a + 14.

5

FEATURES OF PROTEIN ENCODED BY GENE NO: 19

In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: ALVISNPLL (SEQ ID NO:372), PYINTQMCVSSRNKFCISG
 10 HQKYDSHGRETRFEMHKARASSWKNILKIRSLKIISRGFEITNA (SEQ ID NO:373), KFCISGHQKYDSHGRETRFEMHKARAS (SEQ ID NO:374), HTLLEI ANPLQAAVLGASSIHPSIHTSTHLMFMGLKWTELHHSPDSVQGAGAAEAAQTR
 HSLRPGRGRERHDCTLKNLTLFIIC (SEQ ID NO:375), NPLQAAVLGASSIHP
 SIHTSTH (SEQ ID NO:376), and/or SLRPGRGRERHDCTLKN (SEQ ID NO:377).
 15 Polynucleotides encoding these polypeptides are also encompassed by the invention.

This gene is expressed primarily in neutrophils.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are
 20 not limited to, immune or hematopoietic disorders, such as neutropenia, inflammatory, or allergic conditions. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or
 25 lower levels may be routinely detected in certain tissues or cell types (e.g., immune, hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an
 30 individual not having the disorder.

The tissue distribution in neutrophils indicates polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis, treatment, and/or prevention of immune disorders involving neutrophils, or more generally, immune or hematopoietic disorders. Moreover, the expression of this gene product indicates a role in the
 35 regulation of the proliferation; survival; differentiation; and/or activation of hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other

processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's disease, scleroderma and tissues. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:29 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 814 of SEQ ID NO:29, b is an integer of 15 to 828, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:29, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 20

This gene is expressed primarily in 7 week old early stage human.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, fetal or developmental abnormalities, particularly congenital defects, including metabolic conditions. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells,

particularly of the developmental systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., developmental, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, amniotic fluid, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in 7 week old early stage human tissue indicates polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis, treatment, and/or prevention of fetal abnormalities. Expression within embryonic tissue indicates that this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis and treatment of cancer and other proliferative disorders. Similarly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation. Thus this protein may also be involved in apoptosis or tissue differentiation and could again be useful in cancer therapy. Furthermore, the protein is useful in the diagnosis, prevention, and/or treatment of various metabolic disorders such as Tay-Sachs disease, phenylketonuria, galactosemia, hyperlipidemias, porphyrias, leukemias, or Hurler's syndrome. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:30 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 567 of SEQ ID NO:30, b is an integer of 15 to 581, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:30, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 21

In specific embodiments, polypeptides of the invention comprise the following amino acid sequence:
AENVHCTPAWETGRDSEDGKGREGMGRDRKGWDGTGLDGTGWEGKRERNV

PA (SEQ ID NO:378), GRDSEDGKGREGMGRDRKGWDGTGLD (SEQ ID NO:379), TSLGDLWDYNNSSH (SEQ ID NO:380), DRRIIRTREAAVAVSRERP LHSSLGNRERLRRWEGTGRDGKGQEGMGRDGTGWDGGMGREERKKCPS (SEQ ID NO:381), RPLHSSLGNRERLRRWEGTGRDGKG (SEQ ID NO:382),
 5 NQSWGPMGL (SEQ ID NO:383), GGGGCSEPRTSIALQPGKQGETPKMGRD GKGWEGTGRDGTGRDWMGRDGKGREKEMSQQ (SEQ ID NO:384), KQGE TPKMGRDGKGWEGTGRDGTG (SEQ ID NO:385), and/or PVLGTYGTITTPV TELTKGQEKEGGVETVLYE (SEQ ID NO:386). Polynucleotides encoding these polypeptides are also encompassed by the invention.

10 This gene is expressed primarily in frontal cortex from a patient suffering from schizophrenia.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are
 15 not limited to, neural disorders, such as Schizophrenia. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the central nervous system, expression of this gene at significantly higher or lower levels may be routinely detected in certain
 20 tissues or cell types (e.g., neural, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

25 The tissue distribution in frontal cortex tissue indicates polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis, treatment, and/or prevention of some central nervous system disorders, for example, schizophrenia. Moreover, polynucleotides and polypeptides corresponding to this gene are useful for the detection/treatment of neurodegenerative disease states, behavioural disorders, or
 30 inflammatory conditions such as Alzheimers Disease, Parkinsons Disease, Huntingtons Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic disorder, learning
 35 disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, elevated expression of this gene product in regions of the brain indicates that it plays a role in normal neural

function. Potentially, this gene product is involved in synapse formation, neurotransmission, learning, cognition, homeostasis, or neuronal differentiation or survival. Moreover, the gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the developing embryo, sexually-linked disorders, or disorders of the cardiovascular system. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:31 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 775 of SEQ ID NO:31, b is an integer of 15 to 789, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:31, and where b is greater than or equal to a + 14.

20 FEATURES OF PROTEIN ENCODED BY GENE NO: 22

In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: KIVFIDQKWSK (SEQ ID NO:387), CSLFWGILFLSRLRIH LFLSLKPCMCLRPIDILSHFLDIFVTSVLSELEKSSLKTTETFSFAVFLLLMMN (SEQ ID NO:388), LSRLRIHLFLSLKPCMCLRPIDILSH (SEQ ID NO:389), and/or VLSELEKSSLKTTETFSFAVFL (SEQ ID NO:390). Polynucleotides encoding these polypeptides are also encompassed by the invention.

This gene is expressed primarily in thymus and neutrophils.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune or hematopoietic disorders, particularly inflammation, or disorders related to immune cell maturation and/or activation. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain

tissues or cell types (e.g., immune, hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:146 as residues: Lys-38 to Leu-46.

The tissue distribution in thymus and neutrophils indicates polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis, treatment, and/or prevention of inflammatory disorders, such as psoriasis, inflammatory bowel disease, rheumatoid arthritis, and sepsis. Moreover, the expression of this gene product indicates a role in regulating the proliferation; survival; differentiation; and/or activation of hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's disease, scleroderma and tissues. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:32 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome.

Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 870 of SEQ ID NO:32, b is an integer of 15 to

884, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:32, and where b is greater than or equal to a + 14.

5 FEATURES OF PROTEIN ENCODED BY GENE NO: 23

In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: TLFYILH (SEQ ID NO:391). Polynucleotides encoding these polypeptides are also encompassed by the invention.

10 This gene is expressed primarily in bone marrow.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, diseases and disorders afflicting blood cells. Similarly, polypeptides and
15 antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the hematopoietic system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., immune, hematopoietic, and cancerous and wounded tissues) or
20 bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID
25 NO:147 as residues: Pro-30 to Asn-36.

The tissue distribution in bone marrow indicates polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis, treatment, and/or prevention of diseases afflicting the blood, including leukemia, neutropenia, anemia, and stem cell protection during chemotherapy. Moreover, polynucleotides and
30 polypeptides corresponding to this gene are useful for the treatment and diagnosis of hematopoietic related disorders such as anemia, pancytopenia, leukopenia, thrombocytopenia or leukemia since stromal cells are important in the production of cells of hematopoietic lineages. The uses include bone marrow cell ex-vivo culture, bone marrow transplantation, bone marrow reconstitution, radiotherapy or
35 chemotherapy of neoplasia. The gene product may also be involved in lymphopoiesis, therefore, it can be used in immune disorders such as infection, inflammation, allergy, immunodeficiency etc. In addition, this gene product may have commercial utility in the

expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

- 5 Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:33 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome.
- 10 Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 852 of SEQ ID NO:33, b is an integer of 15 to 866, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:33, and where b is greater than or equal to a + 14.

15

FEATURES OF PROTEIN ENCODED BY GENE NO: 24

This gene is expressed primarily in Merkel cells.

- 20 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, integumentary disorders, particularly aberrations in mechanosensory function. Similarly, polypeptides and antibodies directed to these polypeptides are
- 25 useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly in tissues involved in sensory function, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., integumentary, sensory, and cancerous and wounded tissues) or bodily fluids (e.g.,
- 30 lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

- 35 The tissue distribution in Merkel cells indicates polynucleotides and polypeptides corresponding to this gene are useful for Merkel cell dysfunctions, which may include aberrations in sensory function. Alternatively, polynucleotides and polypeptides corresponding to this gene are useful for the treatment, diagnosis, and/or

prevention of various skin disorders including congenital disorders (i.e. nevi, moles, freckles, Mongolian spots, hemangiomas, port-wine syndrome), integumentary tumors (i.e. keratoses, Bowen's disease, basal cell carcinoma, squamous cell carcinoma, malignant melanoma, Paget's disease, mycosis fungoides, and Kaposi's sarcoma),
5 injuries and inflammation of the skin (i.e. wounds, rashes, prickly heat disorder, psoriasis, dermatitis), atherosclerosis, urticaria, eczema, photosensitivity, autoimmune disorders (i.e. lupus erythematosus, vitiligo, dermatomyositis, morphea, scleroderma, pemphigoid, and pemphigus), keloids, striae, erythema, petechiae, purpura, and xanthelasma. In addition, such disorders may predispose increased susceptibility to
10 viral and bacterial infections of the skin (i.e. cold sores, warts, chickenpox, molluscum contagiosum, herpes zoster, boils, cellulitis, erysipelas, impetigo, tinea, athlete's foot, and ringworm). Moreover, the protein product of this gene may also be useful for the treatment or diagnosis of various connective tissue disorders such as arthritis, trauma, tendonitis, chondromalacia and inflammation, autoimmune disorders such as
15 rheumatoid arthritis, lupus, scleroderma, and dermatomyositis as well as dwarfism, spinal deformation, and specific joint abnormalities as well as chondrodysplasias (i.e. spondyloepiphyseal dysplasia congenita, familial osteoarthritis, Atelosteogenesis type II, metaphyseal chondrodysplasia type Schmid). Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets
20 for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:34 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the
25 scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1680 of SEQ ID NO:34, b is an integer of 15 to 1694, where both a and b correspond to the positions of nucleotide residues shown
30 in SEQ ID NO:34, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 25

35 The translation product of this gene shares sequence homology with dihydropyridine receptor or nitrate transporter which are thought to be important in transport of small molecules across the cell membrane. In specific embodiments,

polypeptides of the invention comprise the following amino acid sequence:
GTSFSVLSLIHDTG (SEQ ID NO:392). Polynucleotides encoding these polypeptides
are also encompassed by the invention. The gene encoding the disclosed cDNA is
believed to reside on chromosome 11. Accordingly, polynucleotides related to this
5 invention are useful as a marker in linkage analysis for chromosome 11.

This gene is expressed primarily in kidney cortex and muscle tissue from a
patient with multiple sclerosis, and to a lesser extent, in fetal liver/spleen.

Therefore, polynucleotides and polypeptides of the invention are useful as
reagents for differential identification of the tissue(s) or cell type(s) present in a
10 biological sample and for diagnosis of diseases and conditions which include, but are
not limited to, muscle, urogenital, or renal disorders, particularly musculodegenerative
conditions such as multiple sclerosis, in addition to kidney or metabolic disorders and
diseases. Similarly, polypeptides and antibodies directed to these polypeptides are
useful in providing immunological probes for differential identification of the tissue(s)
15 or cell type(s). For a number of disorders of the above tissues or cells, particularly of
the multiple sclerosis and renal system, expression of this gene at significantly higher or
lower levels may be routinely detected in certain tissues or cell types (e.g., muscle,
urogenital, renal, hepatic, metabolic, immune, hematopoietic, and cancerous and
wounded tissues) or bodily fluids (e.g., lymph, serum, bile, amniotic fluid, plasma,
20 urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an
individual having such a disorder, relative to the standard gene expression level, i.e.,
the expression level in healthy tissue or bodily fluid from an individual not having the
disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID
25 NO:149 as residues: Ala-66 to Leu-73.

The tissue distribution in kidney cortex and muscle tissue, combined with the
homology to small molecule transporters indicates polynucleotides and polypeptides
corresponding to this gene are useful for the diagnosis, treatment, and/or prevention of
disorders of renal functions and muscular diseases, including multiple sclerosis,
30 muscular dystrophy, cardiomyopathy, fibroids, myomas, and rhabdomyosarcomas.
The tissue distribution in kidney indicates that this gene or gene product could be used
in the treatment and/or detection of kidney diseases including renal failure, nephritis,
renal tubular acidosis, proteinuria, pyuria, edema, pyelonephritis, hydronephritis,
nephrotic syndrome, crush syndrome, glomerulonephritis, hematuria, renal colic and
35 kidney stones, in addition to Wilm's Tumor Disease, and congenital kidney
abnormalities such as horseshoe kidney, polycystic kidney, and Falconi's syndrome.
Protein, as well as, antibodies directed against the protein may show utility as a tumor

marker and/or immunotherapy targets for the above listed tissues. The secreted protein can also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions and as nutritional supplements. It may also have a very wide range of biological activities. Typical of these are cytokine, cell proliferation/differentiation modulating activity or induction of other cytokines; immunostimulating/immunosuppressant activities (e.g. for treating human immunodeficiency virus infection, cancer, autoimmune diseases and allergy); regulation of hematopoiesis (e.g. for treating anemia or as adjunct to chemotherapy); stimulation or growth of bone, cartilage, tendons, ligaments and/or nerves (e.g. for treating wounds, stimulation of follicle stimulating hormone (for control of fertility); chemotactic and chemokinetic activities (e.g. for treating infections, tumors); hemostatic or thrombolytic activity (e.g. for treating hemophilia, cardiac infarction etc.); anti-inflammatory activity (e.g. for treating septic shock, Crohn's disease); as antimicrobials; for treating psoriasis or other hyperproliferative diseases; for regulation of metabolism, and behavior. Also contemplated is the use of the corresponding nucleic acid in gene therapy procedures. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. .

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:35 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1201 of SEQ ID NO:35, b is an integer of 15 to 1215, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:35, and where b is greater than or equal to a + 14. .

30

FEATURES OF PROTEIN ENCODED BY GENE NO: 26

In specific embodiments, polypeptides of the invention comprise the following amino acid sequence:

35 VLISASTIGSRTSGAQGMEKMTIPTLAVGEPKTPEKSKCSLKQCFSSCNVHIDH
LGLLLKCKF (SEQ ID NO:393), ASTIGSRTSGAQGMEKMTIPTLA (SEQ ID

NO:394), and/or GEPKTPEKSKCSLKQCFSSCNVHIDHL (SEQ ID NO:395).

Polynucleotides encoding these polypeptides are also encompassed by the invention.

This gene is expressed primarily in kidney medulla.

Therefore, polynucleotides and polypeptides of the invention are useful as
5 reagents for differential identification of the tissue(s) or cell type(s) present in a
biological sample and for diagnosis of diseases and conditions which include, but are
not limited to, renal, urogenital, or more generally, disorders afflicting endothelial
tissues. Similarly, polypeptides and antibodies directed to these polypeptides are useful
10 in providing immunological probes for differential identification of the tissue(s) or cell
type(s). For a number of disorders of the above tissues or cells, particularly of the renal
system, expression of this gene at significantly higher or lower levels may be routinely
detected in certain tissues or cell types (e.g., renal, urogenital, endothelial, and
cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine,
15 synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual
having such a disorder, relative to the standard gene expression level, i.e., the
expression level in healthy tissue or bodily fluid from an individual not having the
disorder.

The tissue distribution in kidney medulla indicates polynucleotides and
polypeptides corresponding to this gene are useful for the diagnosis, treatment, and/or
20 prevention of renal disorders, including lesions, vascular diseases, hydronephrosis,
and renal diseases associated with systemic disorders. Moreover, the gene or gene
product could be used in the treatment and/or detection of kidney diseases including
renal failure, nephritis, renal tubular acidosis, proteinuria, pyuria, edema,
pyelonephritis, hydronephritis, nephrotic syndrome, crush syndrome,
25 glomerulonephritis, hematuria, renal colic and kidney stones, in addition to Wilm's
Tumor Disease, and congenital kidney abnormalities such as horseshoe kidney,
polycystic kidney, and Falconi's syndrome. The protein product can also be used for
the treatment, detection, and/or prevention of various endothelial disorders, which
include microvascular disease, embolism, aneurysm, stroke, or atherosclerosis.
30 Protein, as well as, antibodies directed against the protein may show utility as a tumor
marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available
and accessible through sequence databases. Some of these sequences are related to SEQ
ID NO:36 and may have been publicly available prior to conception of the present
35 invention. Preferably, such related polynucleotides are specifically excluded from the
scope of the present invention. To list every related sequence is cumbersome.
Accordingly, preferably excluded from the present invention are one or more

polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1780 of SEQ ID NO:36, b is an integer of 15 to 1794, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:36, and where b is greater than or equal to a + 14.

5

FEATURES OF PROTEIN ENCODED BY GENE NO: 27

In specific embodiments, polypeptides of the invention comprise the following amino acid sequence:

10 RIRSQDLAIMTTCTFKKYEFSFFVLGFLRRWGATLCLGFTSFAIKFHPSSLCSEKE
GKDFSGFALSIGHPERKKEEGWARWLTVPVVLWEAEVGGSPPEVSS (SEQ ID
NO:396), TTCFKKYEFSFFVLGFLRRWGA (SEQ ID NO:397), SEKEGKDFSGF
ALSIGHPERKKEEGW (SEQ ID NO:398), MNECIAKPCMAAFCSCPSCCLPSR
15 PGCSREQRCAFSCEPCHTVEHWVEPMGQGQRQEHTQGSVLPSSHPSRGKATT
VHSCCQEPWG (SEQ ID NO:399), FCSCPSCCLPSRPGCSREQRCAFSCEP (SEQ
ID NO:400), GQRQEHTQGSVLPSSHPSRGKAT (SEQ ID NO:401), GVVNSCLL
PLPPRLATGMDCGGFASRRMGGRQHAALS VFLPLPLAHGLYPMFNCVAGLT
GKGTSLLSGAARPAGEAAARAGTKGSHARFGNAFIHSF HSFIECLLNTYCV
20 SSALTAVGIGDILKNKNDKSSCLCSC (SEQ ID NO:402), GMDCGGFASRRMG
GRQHAALS VFLP (SEQ ID NO:403), LTGKGTSLLSGAARPAGEAAARAGT (SEQ
ID NO:404), and/or LNTYCVPSALTAVGIGDILKN (SEQ ID NO:405).

Polynucleotides encoding these polypeptides are also encompassed by the invention.

This gene is expressed primarily in fetal lung.

25 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, pulmonary or developmental disorders and/or diseases. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing

30 immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the pulmonary system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., pulmonary, developmental, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, pulmonary surfactant

35 or sputum, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression

level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

5 The tissue distribution in fetal lung indicates polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and intervention of diseases related to pulmonary functions and infections. Moreover, the expression within fetal tissue indicates that this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis and treatment of cancer and other proliferative disorders. Similarly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation. Thus this protein may also be involved in apoptosis or
10 tissue differentiation and could again be useful in cancer therapy. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ
15 ID NO:37 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of
20 a-b, where a is any integer between 1 to 1160 of SEQ ID NO:37, b is an integer of 15 to 1174, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:37, and where b is greater than or equal to a + 14.

25 FEATURES OF PROTEIN ENCODED BY GENE NO: 28

This gene is expressed primarily in hepatocellular tumors.

Therefore, polynucleotides and polypeptides of the invention are useful as
30 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, metabolic or hepatic disorders or diseases, particularly hepatocellular tumors. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the liver,
35 expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., metabolic, hepatic, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, bile, plasma, urine, synovial fluid and

spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in hepatocellular tumors indicates polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and treatment of hepatic disorders, particularly proliferative conditions such as hepatocellular tumors. Moreover, polynucleotides and polypeptides corresponding to this gene are useful for the detection and treatment of liver disorders and cancers (e.g. hepatoblastoma, jaundice, hepatitis, liver metabolic diseases and conditions that are attributable to the differentiation of hepatocyte progenitor cells). In addition, the protein may play a role in the treatment, detection, and/or prevention of developmental abnormalities, fetal deficiencies, pre-natal disorders and various wound-healing models and/or tissue trauma. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:38 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1073 of SEQ ID NO:38, b is an integer of 15 to 1087, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:38, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 29

The translation product of this gene has homology to a contains a helix-loop-helix motif from a *Caenorhabditis elegans* protein (See Genbank Accession No. gi11326280) which is thought to function as a modulator of gene expression. In specific embodiments, polypeptides of the invention comprise the following amino acid sequence:

TSLSQLWHFCHFWPVKFCCGGCPVHCRMFSISGLYLLNASAPSLQLNDPKCL
QT (SEQ ID NO:406), and/or WPVKFCCGGCPVHCRMFSISGLYLLNA (SEQ ID NO:407). Polynucleotides encoding these polypeptides are also encompassed by the invention.

This gene is expressed primarily in normal breast.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, reproductive, or endocrine disorders, particularly of the breast, such as breast cancer. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the cancer and metabolic systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., reproductive, endocrine, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, breast milk, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in breast indicates polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis, treatment, and/or prevention of some types of breast cancer. The protein can also be used for the treatment, detection, and/or prevention of disorders related to ductile tissues or cell types, particularly secretory dysfunctions. The protein can also be used for the treatment of vascular disorders such as atherosclerosis, microvascular disease, embolism, stroke, or aneurysm. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:39 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 424 of SEQ ID NO:39, b is an integer of 15 to 438, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:39, and where b is greater than or equal to a + 14.

35

FEATURES OF PROTEIN ENCODED BY GENE NO: 30

When tested against K562 cell lines, supernatants removed from cells containing this gene activated the ISRE (interferon-sensitive responsive element) promoter element. Thus, it is likely that this gene activates leukemia cells, or potentially
 5 other cells or cell-types, through the JAK-STAT signal transduction pathway. ISRE is a promoter element found upstream in many genes which are involved in the Jak-STAT pathway. The Jak-STAT pathway is a large, signal transduction pathway involved in the differentiation and proliferation of cells. Therefore, activation of the Jak-STAT pathway, reflected by the binding of the ISRE element, can be used to indicate proteins
 10 involved in the proliferation and differentiation of cells. In specific embodiments, polypeptides of the invention comprise the following amino acid sequence:
 SCRCWALGAGGGQRQWVGRS (SEQ ID NO:408), TGAQAPKMGARQKRPL
 QTRIKNSSKSTLWPPQWVRCGRWWTWPSRKTSRPRQLFTSTLSTSASALV
 WPVSWFSQEGH (SEQ ID NO:409), MGARQKRPLQTRIKNSSKSTLWPP (SEQ
 15 ID NO:410), and/or PRRQLFTSTLSTSASALVWPVSW (SEQ ID NO:411).
 Polynucleotides encoding these polypeptides are also encompassed by the invention.

This gene is expressed primarily in human testes.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a
 20 biological sample and for diagnosis of diseases and conditions which include, but are not limited to, reproductive disorders, particularly abnormalities of the testes. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the reproductive
 25 system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., reproductive, testicular, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, seminal fluid, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the
 30 expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:154 as residues: Leu-26 to Glu-52, Gln-71 to Lys-79.

The tissue distribution in testes, combined with the detected ISRE biological
 35 activity, indicates polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis, treatment, and/or prevention of abnormalities of the testes, such as male infertility and proliferative conditions. and/or could be used as a male

contraceptive. The protein can also be used for the maintenance normal testicular function. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available
5 and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:40 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more
10 polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 720 of SEQ ID NO:40, b is an integer of 15 to 734, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:40, and where b is greater than or equal to a + 14.

15

FEATURES OF PROTEIN ENCODED BY GENE NO: 31

This gene is expressed primarily in colon, and to a lesser extent, in thymus.

Therefore, polynucleotides and polypeptides of the invention are useful as
20 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, gastrointestinal, immune, or hematopoietic disorders, particularly abnormalities of the colon, and cancers. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential
25 identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the digestive system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., gastrointestinal, immune, hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal
30 fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in colon and thymus tissue indicates polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis, treatment, and/or
35 prevention of abnormalities of the colon. The protein can also be used for treating inflammatory conditions, or potentially in modulating immune system activation in the treatment of gastrointestinal disorders. Protein, as well as, antibodies directed against

the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:41 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1332 of SEQ ID NO:41, b is an integer of 15 to 1346, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:41, and where b is greater than or equal to a + 14.

15 FEATURES OF PROTEIN ENCODED BY GENE NO: 32

In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: DGGGKEEGVSCLKISLLCGPWLWLP (SEQ ID NO:412), HEMGELAICHTRVPFSLPSSAQGV PQNLQGPIGHLAVCTPSSLTSHWFPQKREKW
20 STVNKRQRFLQFPAPLRNWIPQTPLSLSVSSGPLGSFTVFTLLSLCAWPWCCRD
CYKSCCPIPIFNLTAPLCVHTPEPSS (SEQ ID NO:413), SSAQGV PQNLQGPIGHLAVCTPS (SEQ ID NO:414), VNKRQRFLQFPAPLRNWIPQTPLSLSVS (SEQ ID NO:415), CCRDCYKSCCPIPI FNLTAPLCV (SEQ ID NO:416), DLNVTNEGEGKE
VLGQGSTNNEKKCQKATSNTEPRAREAKARHANMGTS DRESPTWSLTAE
25 GLKAKSKMQGKATKGAASTMGSHNQPHKREIFKHETPSSFPPPSQCQPE
LLPYKYWATLASGYVPSWLP SVDSYRINTAIKDKNGQDT (SEQ ID NO:417),
VLGQGSTNNEKKCQKA TSNTEPRA (SEQ ID NO:418), RESPTWSLTAE
GLKAKSKMQGKATKGAAS (SEQ ID NO:419), and/or GYVPSWLP SVDSYRI
NTAIKDK (SEQ ID NO:420). Polynucleotides encoding these polypeptides are also
30 encompassed by the invention.

This gene is expressed primarily in rhabdomyosarcoma, and to a lesser extent in heart and thymus.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, muscle disorders, particularly rhabdomyosarcoma and other proliferative conditions. Similarly, polypeptides and antibodies directed to these polypeptides are

useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the muscular system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., muscle, and cancerous and
 5 wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID
 10 NO:156 as residues: Gly-28 to Asp-33.

The tissue distribution in rhabdomyosarcoma tissue indicates polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis, treatment, and/or prevention of rhabdomyosarcoma. Moreover, polynucleotides and polypeptides corresponding to this gene are useful for the detection, treatment, and/or prevention of
 15 various muscle disorders, such as muscular dystrophy, cardiomyopathy, fibroids, or myomas. The protein can also be used for the amelioration of proliferative conditions in other tissues, including modulation of the immune response to such tissues. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:42 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome.
 20 Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 984 of SEQ ID NO:42, b is an integer of 15 to 998, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:42, and where b is greater than or equal to a + 14.

30

FEATURES OF PROTEIN ENCODED BY GENE NO: 33

In specific embodiments, polypeptides of the invention comprise the following
 35 amino acid sequence: NSAEQSMILILVT (SEQ ID NO:421), RXDRXPVPELPGYEPT
 RTDISSFKNYRYAFDFARDKDKRSLDIDTAKSMLALLLGRTWPLFSVFYQYLE
 QSKYRVMNKDQWYNVLEFSRTVHADLSNYDEDGAWPVLLDEFVEWQKVRQT

S (SEQ ID NO:422), PTRTDISSFKNIYRYAFDFARDKDQRSL (SEQ ID NO:423), SMLALLLGRTWPLFSVFYQYLE QSKYRVM (SEQ ID NO:424), FSRTVHADLSN YDEDGAWPVLLDEFVE (SEQ ID NO:425), IYRYAFDFAR (SEQ ID NO:426), KD QRSLDI (SEQ ID NO:427), NVLEFSRT (SEQ ID NO:428), and/or DLSNYDEDGA
5 WPVLLDEFVEW (SEQ ID NO:429). Polynucleotides encoding these polypeptides are also encompassed by the invention. The gene encoding the disclosed cDNA is believed to reside on chromosome 11. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 11.

10 This gene is expressed primarily in aortic endothelium, and to a lesser extent, in cancers.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, endothelial disorders, particularly abnormalities of the vascular system
15 and cancers. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the vascular system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., endothelial, vascular, and
20 cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

25 Preferred epitopes include those comprising a sequence shown in SEQ ID NO:157 as residues: Arg-22 to Lys-31.

The tissue distribution in aortic endothelium indicates polynucleotides and polypeptides corresponding to this gene are useful for the treatment, detection, and/or prevention of abnormalities of the vascular system (i.e. embolism, atherosclerosis,
30 aneurysm, stroke, microvascular disease, etc.) and cancers. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ
35 ID NO:43 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome.

Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 644 of SEQ ID NO:43, b is an integer of 15 to 658, where both a and b correspond to the positions of nucleotide residues shown in
5 SEQ ID NO:43, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 34

10 Polypeptides of the invention do not comprise the polypeptide sequence shown as Genbank Accession W59652, which is hereby incorporated herein by reference. In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: LFRCPGKAGTPAGXGPEFPGRPTRPVREKELTETFE (SEQ ID NO:430), FFVFPYPYPFRPLPPIPFPRFPWFRRNFPIPIPESAPTTPLPSEK (SEQ ID
15 NO:432), PWFRRNFPIPIPESAPTTPLP (SEQ ID NO:433), and/or GKAGTPAGXGPEFPGRPTRPV (SEQ ID NO:431). Polynucleotides encoding these polypeptides are also encompassed by the invention.

This gene is expressed primarily in Hodgkin's lymphoma.

Therefore, polynucleotides and polypeptides of the invention are useful as
20 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune or hematopoietic disorders, particularly Hodgkin's lymphoma. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell
25 type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., immune, hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual
30 having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:158 as residues: Ser-21 to Asp-35, Pro-47 to Pro-52, Pro-62 to Asn-67.

35 The tissue distribution in Hodgkin's lymphoma tissue indicates polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis, treatment, and/or prevention of Hodgkin's lymphoma. Moreover, polynucleotides and

polypeptides corresponding to this gene are useful for the treatment and diagnosis of hematopoietic-related disorders such as anemia, pancytopenia, leukopenia, thrombocytopenia or leukemia since stromal cells are important in the production of cells of hematopoietic lineages. The uses include bone marrow cell ex- vivo culture, bone marrow transplantation, bone marrow reconstitution, radiotherapy or chemotherapy of neoplasia. The gene product may also be involved in lymphopoiesis, therefore, it can be used in immune disorders such as infection, inflammation, allergy, immunodeficiency etc. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:44 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 552 of SEQ ID NO:44, b is an integer of 15 to 566, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:44, and where b is greater than or equal to a + 14.

25 FEATURES OF PROTEIN ENCODED BY GENE NO: 35

When tested against U937 and K562 cell lines, supernatants removed from cells containing this gene activated both the GAS (gamma activating sequence), and the ISRE (interferon-sensitive responsive element) promoter elements. Thus, it is likely that this gene activates pro-myeloid, leukemic, or more generally, other cells or cell-types, through the JAK-STAT signal transduction pathway. GAS is a promoter element found upstream of many genes which are involved in the Jak-STAT pathway. The Jak-STAT pathway is a large, signal transduction pathway involved in the differentiation and proliferation of cells. Therefore, activation of the Jak-STAT pathway, reflected by the binding of the GAS element, can be used to indicate proteins involved in the proliferation and differentiation of cells. ISRE is a promoter element found upstream in many genes which are involved in the Jak-STAT pathway. The Jak-STAT pathway is a

large, signal transduction pathway involved in the differentiation and proliferation of cells. Therefore, activation of the Jak-STAT pathway, reflected by the binding of the ISRE element, can be used to indicate proteins involved in the proliferation and differentiation of cells. The translation product of this gene was shown to have

5 homology to a conserved trypsin inhibitor which is thought to play an essential role in protein metabolism and regulation (See Genbank Accession No. pirlS35098|S35098). In specific embodiments, polypeptides of the invention comprise the following amino acid sequence:

10 FYPPMTQGKESLPLLALQIFNTTFRPSFAFFSGHRTLFFGVRSPNPPKPRIFLIW
LIAVAL (SEQ ID NO:434), LLALQIFNTTFRPSFAFFSGHRTLFFGVRSP (SEQ ID
NO:435), HLAQTVMMHPQKSFYQVKNTNHS DRGAIEETQILEDRLGQIPLCLES
QIWEA (SEQ ID NO:436), KNTNHS DRGAIEET QILEDRLGQIPLCL (SEQ ID
NO:437), QGCYRRDS NIGRQVRPDSIMLRKPD LGSITHYGSVLGNLNYCDLP
QLYRNPSLGNSGMREMFSPFYNPVECHP (SEQ ID NO:438), PDSIMLRKPD
15 LGSITHYGSVLGN (SEQ ID NO:439), and/or YRNPSLGNSGMREMFSPFYNPV
(SEQ ID NO:440). Polynucleotides encoding these polypeptides are also encompassed by the invention.

This gene is expressed primarily in brain frontal cortex.

Therefore, polynucleotides and polypeptides of the invention are useful as

20 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, neural disorders, particularly disorders of the central nervous system or endocrine system. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the

25 tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the central nervous system or endocrine system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., neural, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample

30 taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in brain frontal cortex, combined with the detected GAS and ISRE biological activities indicates polynucleotides and polypeptides corresponding

35 to this gene are useful for diagnosis or treatment of disorders of the central nervous system, caused by trauma, inflammation, demyelination, neoplasia, and degenerative diseases. Additionally, the molecule may function as a neuropeptide or hormone.

Moreover, considering the homology to a trypsin inhibitor and its localization in the brain, indicates polynucleotides and polypeptides corresponding to this gene are useful for the detection/treatment of neurodegenerative disease states, behavioral disorders, or inflammatory conditions which include, but are not limited to Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, depression, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, elevated expression of this gene product in regions of the brain indicates that it plays a role in normal neural function. Potentially, this gene product is involved in synapse formation, neurotransmission, learning, cognition, homeostasis, or neuronal differentiation or survival. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:45 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1263 of SEQ ID NO:45, b is an integer of 15 to 1277, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:45, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 36

30

The translation product of this gene was found to have homology to a zinc finger protein from *Mus musculus* (See Genbank Accession No. gnllPIDle225687) which is thought to be involved in the modulation of gene regulation. In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: NSARGLSGGHPFPWLSEGHFP (SEQ ID NO:441), TSDSLTLGILLGI YTNHIWEMFLAASRINSPKLEPEKSVKRQINFSSKDVGCSSLEVPKDGPP LSHGKEWIPLSHRKGWIPLSHMKGWPSLSHGKGWPP LSPRAEF (SEQ ID

35

NO:442), LGILLLLGIYTNHIWEMFLAA (SEQ ID NO:443), KSVKRQINFPSKDV
GCSLEVPKDGGP (SEQ ID NO:444), GKEWIPLSHRKGWIPLSHMKGWPSLSH
(SEQ ID NO:445), GWASTQPRERMDPAQPQERMDPSQPHERMAL TQPWK RMAP
TQPSCRI (SEQ ID NO:446), and/or PAQPQERMDPSQPHERMAL TQPWK (SEQ ID
5 NO:447). Polynucleotides encoding these polypeptides are also encompassed by the
invention.

This gene is expressed primarily in neutrophils.

Therefore, polynucleotides and polypeptides of the invention are useful as
reagents for differential identification of the tissue(s) or cell type(s) present in a
10 biological sample and for diagnosis of diseases and conditions which include, but are
not limited to, immune or hematopoietic disorders. Similarly, polypeptides and
antibodies directed to these polypeptides are useful in providing immunological probes
for differential identification of the tissue(s) or cell type(s). For a number of disorders
of the above tissues or cells, particularly of the immune system, expression of this gene
15 at significantly higher or lower levels may be routinely detected in certain tissues or cell
types (e.g., immune, hematopoietic cancerous and wounded tissues) or bodily fluids
(e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or
cell sample taken from an individual having such a disorder, relative to the standard
gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an
20 individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID
NO:160 as residues: Ser-30 to Asp-39.

The tissue distribution in neutrophils indicates polynucleotides and polypeptides
corresponding to this gene are useful for the diagnosis, treatment, and/or prevention of
25 immune disorders involving neutrophils, including neutropenia. The expression of this
gene product indicates a role in regulating the proliferation; survival; differentiation;
and/or activation of hematopoietic cell lineages, including blood stem cells. This gene
product may be involved in the regulation of cytokine production, antigen presentation,
or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by
30 boosting immune responses). Since the gene is expressed in cells of lymphoid origin,
the natural gene product may be involved in immune functions. Therefore it may be also
used as an agent for immunological disorders including arthritis, asthma,
immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis,
granulomatous disease, inflammatory bowel disease, sepsis, acne, neutropenia,
35 neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune
reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-
host diseases. or autoimmunity disorders, such as autoimmune infertility, lense tissue

injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's disease, scleroderma and tissues. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. The protein product of this gene can be used in the preparation of therapeutic compositions, for treating, preventing or delaying the recurrence of a tumour or neuronal disorders, e.g. genetic diseases or acquired degenerative encephalopathies such as Alzheimer's disease. Moreover, the protein is also useful in the induction or inhibition of cellular apoptosis resulting in inhibition of tumour cell growth, to suppress tumour formation, to induce G1 arrest of the cell cycle and to act as nuclear transcription factor. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:46 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 428 of SEQ ID NO:46, b is an integer of 15 to 442, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:46, and where b is greater than or equal to a + 14.

25

FEATURES OF PROTEIN ENCODED BY GENE NO: 37

When tested against U937 and K562 cell lines, supernatants removed from cells containing this gene activated both the GAS (gamma activating sequence), and the ISRE (interferon-sensitive responsive element) promoter elements. Thus, it is likely that this gene activates pro-myeloid, leukemic, or more generally, other cells or cell-types, through the JAK-STAT signal transduction pathway. GAS is a promoter element found upstream of many genes which are involved in the Jak-STAT pathway. The Jak-STAT pathway is a large, signal transduction pathway involved in the differentiation and proliferation of cells. Therefore, activation of the Jak-STAT pathway, reflected by the binding of the GAS element, can be used to indicate proteins involved in the proliferation and differentiation of cells. ISRE is a promoter element found upstream in

many genes which are involved in the Jak-STAT pathway. The Jak-STAT pathway is a large, signal transduction pathway involved in the differentiation and proliferation of cells. Therefore, activation of the Jak-STAT pathway, reflected by the binding of the ISRE element, can be used to indicate proteins involved in the proliferation and differentiation of cells. The protein product of this gene was found to have homology to the G-protein coupled receptor TM1 long consensus polypeptide (See Genbank Accession No. R50790) which indicates the protein is useful in the modulation of signalling events, cell-cycle regulation and/or transcriptional regulation. In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: IANGGGRPIKLNALYK IQNECKIVFTCIDF (SEQ ID NO:448), and/or MPCIK SKMNAKLFSLVLTCCMIPISVLFGTCT (SEQ ID NO:449). Polynucleotides encoding these polypeptides are also encompassed by the invention. The gene encoding the disclosed cDNA is believed to reside on chromosome 1. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 1.

This gene is expressed primarily in duodenum.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, gastrointestinal disorders, particularly abnormalities of the duodenum. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the digestive system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., gastrointestinal, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in duodenum, the homology to the TM1 g-protein coupled receptor consensus sequence, in addition to the detected GAS and ISRE biological activities, indicates polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis, treatment, and/or prevention of the abnormalities of the duodenum, particularly proliferative conditions such as cancers. Moreover, the protein can be used in G-protein coupled receptor ligand binding assays. The assay can be used to identify fragments from GPR proteins (see Genseq Accession Nos. R48686-R48758 for examples) which retain biological activity such as binding a GPR ligand or

modulating GPR ligand binding to a GPR (see Genseq Accession Nos. R48759-R48758, R50569-R50807 and R89189-R89195 for examples of polypeptide fragments). The polypeptide fragments can be used in compositions for treating subjects suffering from a pathology related to a GPR abnormality e.g. a psychotic disorder such as schizophrenia. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:47 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 876 of SEQ ID NO:47, b is an integer of 15 to 890, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:47, and where b is greater than or equal to a + 14.

20 FEATURES OF PROTEIN ENCODED BY GENE NO: 38

The translation product of this gene shares sequence homology with a growth and transformation dependent protein (>gil207250), which is thought to be important in the regulation of cellular growth and proliferation. In specific embodiments, polypeptides of the invention comprise the following amino acid sequence:

QVAMGSLSGLRLAAGSCFRLCERDVSSSLRLTRSSDLKRINGFCTKPQESPG
APSRTYNRVPLHKPTDWQKKILIWSGRFKKEDEIPETVSLEMLDAAKNK (SEQ
ID NO:450), GLRLAAGSCFRLCERDVSSSLRLTR (SEQ ID NO:451), APSRTYNR
VPLHKPTDWQKK (SEQ ID NO:452), IWSGRFKKEDEIPETVSLEMLDA (SEQ ID
NO:453), MDFAQNHRKVPPELHPALTTECLYTNLRIGRKRSSYGQVASKRKM
KSQRLSRWRCLMLQTRCE (SEQ ID NO:454), KVPPELHPALTTECLYTNLR
(SEQ ID NO:455), KRSSYGQVASKRKMKSQRLSRWRCLM (SEQ ID NO:456),
INGFCTKPQESP (SEQ ID NO:457), RVPLHKPTD (SEQ ID NO:458), WSGRFK
KE (SEQ ID NO:459), EMLDAAKNK (SEQ ID NO:460), SYLMIALTV (SEQ ID
NO:461), and/or MVIEGKKAA (SEQ ID NO:462). Polynucleotides encoding these
polypeptides are also encompassed by the invention.

This gene is expressed primarily in ovary.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, reproductive, or endocrine disorders, particularly abnormalities of the ovary. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the reproductive system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., reproductive, endocrine, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:162 as residues: Lys-25 to Thr-33, Leu-39 to Glu-47.

The tissue distribution in ovary, combined with the homology to the growth and transformation dependent protein, indicates polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis, treatment, and/or prevention of the abnormalities of the ovary such as ovarian cancer. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:48 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 723 of SEQ ID NO:48, b is an integer of 15 to 737, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:48, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 39

In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: RPGMRALGSCLSLALCSPQARPGPRTLDASTATLTPHF
 5 SPCARFSPVGPSAVPFAATPLPLAGPHQP (SEQ ID NO:463), GSCLSLALCS
 PQARPGPRT (SEQ ID NO:464), HFSPCARFSPVGPSAVPFAATPL (SEQ ID
 NO:465), AIEERNKSRLTQQASEPTGSPRYLHEQHPGSRSQMDCGSLTMXCPPP
 RVRDDRTSARGVPRQAAPDIVGGRPSSRACVSPACAPSAAVFPY (SEQ ID
 NO:466), LTQQASEPTGSPRYLHEQHPGSR (SEQ ID NO:467), and/or SARG
 10 VPRQAAPDIVGGRPSSRACVS (SEQ ID NO:468). Polynucleotides encoding these
 polypeptides are also encompassed by the invention.

This gene is expressed primarily in ovarian tumor.

Therefore, polynucleotides and polypeptides of the invention are useful as
 reagents for differential identification of the tissue(s) or cell type(s) present in a
 15 biological sample and for diagnosis of diseases and conditions which include, but are
 not limited to, reproductive or endocrine disorders, particularly ovarian tumors.
 Similarly, polypeptides and antibodies directed to these polypeptides are useful in
 providing immunological probes for differential identification of the tissue(s) or cell
 type(s). For a number of disorders of the above tissues or cells, particularly of the
 20 female reproductive system, expression of this gene at significantly higher or lower
 levels may be routinely detected in certain tissues or cell types (e.g., reproductive,
 endocrine cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum,
 plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken
 from an individual having such a disorder, relative to the standard gene expression
 25 level, i.e., the expression level in healthy tissue or bodily fluid from an individual not
 having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID
 NO:163 as residues: Met-1 to Gly-6, Trp-23 to Arg-29, Ala-38 to Ser-45.

The tissue distribution in ovarian tumor tissue indicates polynucleotides and
 30 polypeptides corresponding to this gene are useful for the diagnosis, treatment, and/or
 prevention of reproductive disorders, particularly ovarian conditions, such as tumors.
 Protein, as well as, antibodies directed against the protein may show utility as a tumor
 marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available
 35 and accessible through sequence databases. Some of these sequences are related to SEQ
 ID NO:49 and may have been publicly available prior to conception of the present
 invention. Preferably, such related polynucleotides are specifically excluded from the

scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 557 of SEQ ID NO:49, b is an integer of 15 to 571, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:49, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 40

10

In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: PRVRKTPHLSASGK (SEQ ID NO:469), YYYSMLKICHITI LETLSDRTPRKFAK KCYILYIKLSDSSVEKVAYTLLLLIPAAIEKK (SEQ ID NO:470), and/or TILETLSDRTPRKFAK KCYILYIKLSDSSVEK (SEQ ID NO:471).

15 Polynucleotides encoding these polypeptides are also encompassed by the invention.

This gene is expressed primarily in endometrial stromal cells treated with estradiol.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, reproductive disorders, particularly cancer of the endometrium. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the reproductive system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., reproductive, endometrial, and cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:164 as residues: Met-1 to Ser-7.

The tissue distribution in endometrial stromal cells indicates polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis, treatment, and/or prevention of diseases of the endometrium, particularly cancer or diseases caused by hormonal imbalances. Protein, as well as, antibodies directed against the protein may

show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:50 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 342 of SEQ ID NO:50, b is an integer of 15 to 356, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:50, and where b is greater than or equal to a + 14.

15 FEATURES OF PROTEIN ENCODED BY GENE NO: 41

The translation product of this gene shares sequence homology with the smaller hepatocellular oncoprotein which is thought to be important in protein synthesis leading to cellular transformation. In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: VHTKEIFRERSAGFPVK (SEQ ID NO:472), LEMGFQPTKEINARGSEPCQAQSTSLPKLPRWGSRPEAPQTPQGG LESRCCTPVSKQSLNLKADRFKALTLGRAQWLT PVIQALSELRWVDHLRSGV (SEQ ID NO:473), FQPT KEINARGSEPCQAQSTSLPK (SEQ ID NO:474), PKLPR WGSRPEAPQTPQGGLESRCCTP (SEQ ID NO:475), and/or RFKALTLGRAQWLT PVIQALSELRWVD (SEQ ID NO:476). Polynucleotides encoding these polypeptides are also encompassed by the invention.

This gene is expressed primarily in human bladder.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, urogenital disorders, particularly proliferative conditions, such as bladder tumors. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the bladder, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., urogenital, bladder, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid and spinal

fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID
5 NO:165 as residues: Pro-30 to Lys-38, Pro-45 to Ile-60, Leu-79 to Ser-96, His-98 to Gly-118.

The tissue distribution in bladder tumors indicates polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis of carcinomas and preneoplastic or pathological conditions of bladder, or of the urogenital/renal system.

10 Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:51 and may have been publicly available prior to conception of the present
15 invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 899 of SEQ ID NO:51, b is an integer of 15 to
20 913, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:51, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 42

25

In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: RIPLQSDGSFLHEKSSQQRSNRNFPCPTLQCNPEVSFWFV
VTDP SKNHTLPAVEVQSAIRMNKNRINNAFFLNDQTLEFLKIPSTLAPPM DPS
VPIWIIIFGVIFCIIHVAIALLLSGIWQRRRKNKEPSEVDDAEDKCENMITIENGIP
30 SDPLDMKG GHINDAFMTEDERLTPL (SEQ ID NO:477), PCPTLQCNPEVSF
WFVVTDP SKNHT (SEQ ID NO:478), AIRMNKNRINNAFFLNDQTLEFL (SEQ ID
NO:479), IWQRRRKNKEPSEVDDAEDKCENM (SEQ ID NO:480), PLDMKG
GHINDAFMTEDER (SEQ ID NO:481), GSRTTALQRGVSLSSVMKASLICPP
FMSRGSEGMPFSIVIMFSHLSSASSTSDGSLFLLRCQIPDKISSAIATMM
35 MQNITPNIIQMGTDGSMGGASVEGIFKNSRVWSFRKKALLIRFLFILMADCTST
A GRV (SEQ ID NO:482), VSLSSVMKASLICPPFMSRGSEGMPFS (SEQ ID

NO:483), and/or SMGGASVEGIFKNSRVWSFRKKAL (SEQ ID NO:484).

Polynucleotides encoding these polypeptides are also encompassed by the invention.

This gene is expressed primarily in kidney, and to a lesser extent, in gall bladder and testes.

5 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, diseases of the renal, urogenital, or reproductive system. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing
10 immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the renal system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., renal, urogenital, reproductive, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, seminal fluid, synovial
15 fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:166 as residues: Lys-60 to Ala-66, Thr-78 to Ser-83.

20 The tissue distribution in kidney indicates the protein product of this gene could be used in the treatment and/or detection of kidney diseases including renal failure, nephritis, renal tubular acidosis, proteinuria, pyuria, edema, pyelonephritis, hydronephritis, nephrotic syndrome, crush syndrome, glomerulonephritis, hematuria, renal colic and kidney stones, in addition to Wilm's Tumor Disease, and congenital
25 kidney abnormalities such as horseshoe kidney, polycystic kidney, and Falconi's syndrome. Moreover, the tissue distribution in gall bladder indicates that the protein is useful for the treatment, detection, and/or prevention of various metabolic disorders. Alternatively, the expression within testes indicates that the protein is useful in normal testicular function. Therefore, this gene product may be useful in the treatment of male
30 infertility, and/or could be used as a male contraceptive. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ
35 ID NO:52 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome.

Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1342 of SEQ ID NO:52, b is an integer of 15 to 1356, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:52, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 43

10 In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: GARGSQQDAPALQEA EVRGPERAQPARGR (SEQ ID NO:485), SERPGEGPARPGQDDQGP VPAVAGAGVGVHDPADHRVLGQRSAA HFYLHTSFSRPHTGPPLTPGPDRTGSSRPTPMSTSFWTISHAGVKQSDLP RKE TEQPPAPGEHGGERERLRLVPARRPAQPRPGPAAGGAEERAAGLLRQLQP

15 GLPHQGARIRRH PQLGAEPDRGRPARGHLLRAQGGLHQLEARD DRAER KPAAPRCALPRPAAH PARARAQRQ RAPDLQQVLAPLREALPPHEGQAQEVHQ VPLRARPLRAPDLRLPQQVRAGERGVLPQVRR AHAAGVRQPHQPARLGAR GLPRWPQGVLRLQHPVPAGPAHGEAGALQ RALAAGVPPLPPVPDRLRFLG KLETLD EDA AQLLQLLQVDRQSASPRATGTGPPAAGRRTGSPRSPWPGG

20 SSCINSTRPTLFSSATPSPKTSSETESFRVAFSRVPGT (SEQ ID NO:486), RPGQ DDQGP VPAVAGAGVGVH DPA (SEQ ID NO:487), SRPHTGPPLTPGPDRT GSSR (SEQ ID NO:488), SHAGVKQSDLP RKETEQPPAPGE (SEQ ID NO:489), RRPAQPRPGPAAGGAEERAAGLL (SEQ ID NO:490), RRHPQLGAEPDRGR PARGHLL (SEQ ID NO:491), RDDRAERKPAAPRCALPRPAAH PAR (SEQ ID NO:492), RAPDLQQVLAPLREALPPHEGQAQEV (SEQ ID NO:493), DLRLPQQ VRAGERGVLPQVRR AHAAG (SEQ ID NO:494), QPARLGARGLPRWPQGVL RQLHPVPAG (SEQ ID NO:495), AGVPPLPPVPDRLRFLGKLET LDE (SEQ ID NO:496), QLLQLLQVDRQSASPRATGTGPPAA (SEQ ID NO:497), NSTRPTLFSS ATPSPKTSSETESFR (SEQ ID NO:498), LGGKRTAGPPGVAAAAARRPRPE

30 SPASPGIVVDLARVAEAVHLPPVLVEGRQLLRVRVQQVLDEVGEGHLEASA EGLARRGGQAGVVG VHPQHGHGELAVELLVLQLELAAEGGDQAHEGVAHEE ELGVLLELDLHEVAGELPVAAPELVEGQVRAGVVHVLARDAQRVAVGRTA VQQASAQHDH HALPVGAGHLGHVAVDGPVPVVDQVAQLRVGDVVECALLG GEGQAGVGAEAPQHVPPLRLLPALVWAAPGVARGPVVASHALLHAPPA

35 QAAAPSPFWEGHSASRQHEKLSRNSSTSES AVSS LSCPARAWAAAAPCAA (SEQ ID NO:499), EAVHLPPVLVEGRQLLRVRVQQV (SEQ ID NO:500), GHLEA SAEGLARRGGQAGVVG VHP (SEQ ID NO:501), QLELAAEGGDQAHEGVAHE

EELGVLLEL (SEQ ID NO:502), GELPVAAPELVEGQVRAGVVHVLARDA (SEQ ID NO:503), AVQQASAQHDHHPVGAGHLGHVA (SEQ ID NO:504), ALVW AAPGVARGPVVASHALLHA (SEQ ID NO:506), HDQVAQLRVGDVVECALLG GEGQAG (SEQ ID NO:505), PPAQAAAPSPFWEGHSASRQHEKLSRNS (SEQ ID NO:507), SRVTFFERRRSSRLRRGSMEESVRGYDWSRPDARRSPDQGRQQA
 5 RRNVLRGFCANSSLAFTKERAFDDIPNSELSHLIVDDRHGAIYCYVPKV ACTNWKRVMIVLSGSLHRGAPYRDPLRIPREHVHNASAHLTFNKFWRRYGK LSRHLMKVKLKKYTKFLFVRDPFVRLISAFRSKFELENEEFYRKFAVPMLRVY ANHTSLPASAREAFRAGLKVSFANFIQYLLDPHTEKLAPFNEHWRQVYRLC
 10 HPCQIDYDSWGSWRLWTRTPRSCCSYSRWGTGSPLPELPEQDRQQLGGGLVR QD PPGLEAAAV (SEQ ID NO:508), RSPDQGRQQAERRNVLRGFCANSSLA (SEQ ID NO:509), TKERAFDDIPNSELSHLIVDDRHGAIYC (SEQ ID NO:510), FNKFWRRYGKLSRHLMKVKLKKY (SEQ ID NO:511), FVRLISAFRSKFELE NEEFYRKFA (SEQ ID NO:512), TSLPASAREAFRAGLKVSFANFIQYL (SEQ ID NO:513), and/or SYSRWGTGSPLPELPEQDRQQLGGG (SEQ ID NO:514).

Polynucleotides encoding these polypeptides are also encompassed by the invention.

The gene encoding the disclosed cDNA is believed to reside on chromosome 7.

Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 7.

20 It has been discovered that this gene is expressed primarily in PMA activated monocytic HL60 cells.

Therefore, nucleic acids of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of the following diseases and conditions: blood related disease such as
 25 leukemia. Similarly, polypeptides and antibodies directed to those polypeptides are useful to provide immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be detected in certain tissues (e.g., immune, hematopoietic, and cancerous and wounded
 30 tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID
 35 NO:167 as residues: Ala-29 to Thr-37, Pro-39 to Leu-63.

The tissue distribution in HL60 cells suggests the protein product of this clone is useful for the diagnosis, treatment, and/or prevention of blood related diseases such

as leukemia. Moreover, the protein product of this clone is useful for the treatment and diagnosis of hematopoietic related disorders such as anemia, pancytopenia, leukopenia, thrombocytopenia or leukemia since stromal cells are important in the production of cells of hematopoietic lineages. The uses include bone marrow cell ex-vivo culture,
5 bone marrow transplantation, bone marrow reconstitution, radiotherapy or chemotherapy of neoplasia. The gene product may also be involved in lymphopoiesis, therefore, it can be used in immune disorders such as infection, inflammation, allergy, immunodeficiency etc. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the
10 differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ
15 ID NO:53 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of
20 a-b, where a is any integer between 1 to 1533 of SEQ ID NO:53, b is an integer of 15 to 1547, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:53, and where b is greater than or equal to a + 14.

25 FEATURES OF PROTEIN ENCODED BY GENE NO: 44

When tested against fibroblast cell lines, supernatants removed from cells containing this gene activated the EGR1 (early growth response gene 1) promoter element. Thus, it is likely that this gene activates fibroblasts, or more generally, other
30 cells or cell types, through the EGR1 signal transduction pathway. EGR1 is a separate signal transduction pathway from Jak-STAT, genes containing the EGR1 promoter are induced in various tissues and cell types upon activation, leading the cells to undergo differentiation and proliferation. The gene encoding the disclosed cDNA is believed to reside on chromosome 12. Accordingly, polynucleotides related to this invention are
35 useful as a marker in linkage analysis for chromosome 12. In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: STGCSE (SEQ ID NO:515),

CLCLGCGLPELHSYLDPGPYLLVYPTLFWLCPSAVSPWAYTCYQLGLGPQWGA
 AALSFTVDAAIRVWDVSTETCVPLPWFRGGGVNTCSGPQTAAKSWLPLLQLSF
 ESGRPRCGLVRGGLLYQGAVRLAA GAQMAADCCSLYWESH (SEQ ID
 NO:516), YPTLFWLCPSAVSPWAYTCYQLGLGP (SEQ ID NO:517), DVSTETCVP
 5 LPWFRGGGVNTCSGPQ (SEQ ID NO:518), LLYQGAVRLAA GAQMAADCCSL
 (SEQ ID NO:519), NKRKTYLFLEVGMWGVGQNRWWPWVERVPRGRGWGCL
 SKEGQVMNRASTPSRGFLGPPKHWAKTWKLGIDKVQRDVGNSACGPAH
 TEQGPFVEGRWKVMSWGWAPGSPWIMPQGRSSNTGLFRVRKRRMTGLPS
 CTLGFPFISTARRSPLGSQTME (SEQ ID NO:520), GVGQNRWWPWVERVPRG
 10 RGWGCLSKEG (SEQ ID NO:521), AKTWKLGIDKVQRDVGNSACGPAHTE
 (SEQ ID NO:522), and/or WAPGSPWIMPQGRSSNTGLFRVRKRRMTGLPSC
 TLGFPFIST (SEQ ID NO:523). Polynucleotides encoding these polypeptides are also
 encompassed by the invention.

This gene is expressed primarily in fetal tissues such as fetal brain, fetal liver,
 15 fetal kidney, and to a lesser extent, in T cells and macrophages.

Therefore, polynucleotides and polypeptides of the invention are useful as
 reagents for differential identification of the tissue(s) or cell type(s) present in a
 biological sample and for diagnosis of diseases and conditions which include, but are
 not limited to, blood-related, immuno-related, neural-related, or developmental
 20 disorders. Similarly, polypeptides and antibodies directed to these polypeptides are
 useful in providing immunological probes for differential identification of the tissue(s)
 or cell type(s). For a number of disorders of the above tissues or cells, particularly of
 the hematopoiesis and immune system, expression of this gene at significantly higher or
 lower levels may be routinely detected in certain tissues or cell types (e.g., neural,
 25 immune, hematopoietic, urogenital, renal, hepatic, metabolic, developmental, and
 cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, amniotic fluid,
 bile, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from
 an individual having such a disorder, relative to the standard gene expression level, i.e.,
 the expression level in healthy tissue or bodily fluid from an individual not having the
 30 disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID
 NO:168 as residues: Cys-126 to Thr-138, Glu-165 to Gly-172, Thr-189 to Leu-200,
 Gly-222 to Gly-229, Pro-346 to Lys-354.

The tissue distribution in fetal liver indicates polynucleotides and polypeptides
 35 corresponding to this gene are useful for the diagnosis, treatment, and/or prevention of
 blood related diseases, particularly immune or hematopoietic disorders. Alternatively,
 the expression within fetal brain indicates polynucleotides and polypeptides

corresponding to this gene are useful for the detection/treatment of neurodegenerative disease states, behavioral disorders, or inflammatory conditions which include, but are not limited to Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral

5 neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, depression, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, elevated expression of this

10 gene product in regions of the brain indicates that it plays a role in normal neural function. Potentially, this gene product is involved in synapse formation, neurotransmission, learning, cognition, homeostasis, or neuronal differentiation or survival. Alternatively, the expression within fetal kidney indicates the protein product of this gene could be used in the treatment and/or detection of kidney diseases including

15 renal failure, nephritis, renal tubular acidosis, proteinuria, pyuria, edema, pyelonephritis, hydronephritis, nephrotic syndrome, crush syndrome, glomerulonephritis, hematuria, renal colic and kidney stones, in addition to Wilm's Tumor Disease, and congenital kidney abnormalities such as horseshoe kidney, polycystic kidney, and Falconi's syndrome. Moreover, the expression within various

20 fetal tissues, combined with the detected EGR1 biological activity, indicates that this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis and treatment of cancer and other proliferative disorders. Similarly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation. Thus this protein may also be involved in apoptosis or tissue

25 differentiation and could again be useful in cancer therapy. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ

30 ID NO:54 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of

35 a-b, where a is any integer between 1 to 1324 of SEQ ID NO:54, b is an integer of 15 to 1338, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:54, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 45

In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: SSYQCPKVTFFKSSVDT (SEQ ID NO:524). Polynucleotides
5 encoding these polypeptides are also encompassed by the invention. The gene encoding the disclosed cDNA is believed to reside on chromosome 15. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 15.

This gene is expressed primarily in glioblastoma, liver, fetal lung, and
10 amygdala.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, neural, metabolic, or developmental disorders, particularly mental or
15 neurodegenerative conditions. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the the central nervous system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g.,
20 neural, metabolic, developmental, pulmonary, hepatic, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, amniotic fluid, pulmonart surfactant or sputum, bile, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an
25 individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:169 as residues: Pro-31 to Ala-37, Lys-62 to Asn-72.

The tissue distribution in glioblastoma and amygdala indicates polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis, treatment,
30 and/or prevention of central nervous system disorders. Moreover, polynucleotides and polypeptides corresponding to this gene are useful for the detection/treatment of neurodegenerative disease states, behavioral disorders, or inflammatory conditions which include, but are not limited to Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating
35 diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, depression, panic disorder,

learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, elevated expression of this gene product in regions of the brain indicates that it plays a role in normal neural function. Potentially, this gene product is involved in synapse formation, neurotransmission, learning, cognition, homeostasis, or neuronal differentiation or survival. The protein may also be useful in the treatment, detection, and/or prevention of liver disorders, which include, but are not limited to hepatoblastoma, jaundice, hepatitis, liver metabolic diseases and conditions that are attributable to the differentiation of hepatocyte progenitor cells. In addition the expression in fetus would suggest a useful role for the protein product in developmental abnormalities, fetal deficiencies, pre-natal disorders and various wound-healing models and/or tissue trauma. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:55 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 2057 of SEQ ID NO:55, b is an integer of 15 to 2071, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:55, and where b is greater than or equal to a + 14.

25

FEATURES OF PROTEIN ENCODED BY GENE NO: 46

In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: YIYSYLGFNQINK (SEQ ID NO:525). Polynucleotides encoding these polypeptides are also encompassed by the invention.

This gene is expressed in only T-cell helper II cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune or hematopoietic disorders, particularly infectious diseases, inflammatory, or immunodeficiency conditions. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for

differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., immune, hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:170 as residues: Pro-44 to Tyr-49.

The tissue distribution in T-helper cells indicates polynucleotides and polypeptides corresponding to this gene are useful for the treatment of infectious diseases. Moreover, the expression of this gene product indicates a role in regulating the proliferation; survival; differentiation; and/or activation of hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's disease, scleroderma and tissues. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:56 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more

polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1885 of SEQ ID NO:56, b is an integer of 15 to 1899, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:56, and where b is greater than or equal to a + 14.

5

FEATURES OF PROTEIN ENCODED BY GENE NO: 47

The translation product of this gene has been shown to have homology to the human nuclear factor IV (See Genbank Accession No. gi35038), which is thought to play a role as a type 2 DNA helicase in DNA metabolism either during transcription, DNA repair, and/or during the cell-cycle. Moreover, the protein may play a role in chromosomal translocations. In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: ARDLIL (SEQ ID NO:526), LTFYL QFLAPKDKPSGDTAAVFEEGGDVDDLSTFNMHLVFCD (SEQ ID NO:527), and/or FLAPKDKPSGDTAAVFEEGGDVDDL (SEQ ID NO:528). Polynucleotides encoding these polypeptides are also encompassed by the invention. The gene encoding the disclosed cDNA is believed to reside on chromosome 2. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 2.

This gene is expressed primarily in activate T-cells, and to a lesser extent, in B-cells and monocytes.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune or hematopoietic disorders, particularly leukemia, Grave's disease, rheumatoid arthritis and other autoimmune diseases. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., immune, hematopoietic cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:171 as residues: Gly-27 to Cys-35.

The tissue distribution in T-cells, B-cells, and monocytes indicates polynucleotides and polypeptides corresponding to this gene are useful for the treatment or diagnosis of immune system diseases. Moreover, the expression of this gene product indicates a role in regulating the proliferation; survival; differentiation; and/or activation of hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's disease, scleroderma and tissues. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:57 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1529 of SEQ ID NO:57, b is an integer of 15 to 1543, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:57, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 48

In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: ARAGAKILFEGEF (SEQ ID NO:529), NFEIHSAPFMLFVA
 5 CLLHSSCPRTARFLASPLSESNVIFYQNQYQFPCILCFIEFARLTSFKHLHSQSH
 LVRLQYEDFSVSSE AWDTELT (SEQ ID NO:530), FPFMLFVACLLHSSCPRTA
 RFLASPL (SEQ ID NO:531), NVIFYQNQYQFPCILCFIEFARLTSF (SEQ ID
 NO:532), and/or SQSHLVRLQYEDFSVSSE AWDTE (SEQ ID NO:533).

Polynucleotides encoding these polypeptides are also encompassed by the invention.

- 10 The gene encoding the disclosed cDNA is believed to reside on chromosome 14.
 Accordingly, polynucleotides related to this invention are useful as a marker in linkage
 analysis for chromosome 14.

This gene is expressed primarily in fetal tissues such as fetal liver, fetal brain,
 fetal lung and fetal spleen.

- 15 Therefore, polynucleotides and polypeptides of the invention are useful as
 reagents for differential identification of the tissue(s) or cell type(s) present in a
 biological sample and for diagnosis of diseases and conditions which include, but are
 not limited to, developmental disorders and cancers. Similarly, polypeptides and
 antibodies directed to these polypeptides are useful in providing immunological probes
 20 for differential identification of the tissue(s) or cell type(s). For a number of disorders
 of the above tissues or cells, particularly of the immune system and nervous system,
 expression of this gene at significantly higher or lower levels may be routinely detected
 in certain tissues or cell types (e.g., developmental, hepatic, immune, hemaopoietic,
 neural, pulmonary, and cancerous and wounded tissues) or bodily fluids (e.g., serum,
 25 plasma, amniotic fluid, urine, synovial fluid and spinal fluid) or another tissue or cell
 sample taken from an individual having such a disorder, relative to the standard gene
 expression level, i.e., the expression level in healthy tissue or bodily fluid from an
 individual not having the disorder.

- Preferred epitopes include those comprising a sequence shown in SEQ ID
 30 NO:172 as residues: Gly-37 to Asp-46, Ser-48 to Val-54.

- The tissue distribution in fetal tissues indicates polynucleotides and polypeptides
 corresponding to this gene are useful for the diagnosis, treatment, and/or prevention of
 developmental disorders and cancers. Moreover, the expression within embryonic
 tissue and other cellular sources marked by proliferating cells indicates that this protein
 35 may play a role in the regulation of cellular division, and may show utility in the
 diagnosis and treatment of cancer and other proliferative disorders. The protein is also
 useful in the treatment, detection, and/or prevention of immune, hematopoietic,

pulmonary, or metabolic diseases, disorders, and/or conditions. Similarly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation. Thus this protein may also be involved in apoptosis or tissue differentiation and could again be useful in cancer therapy. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:58 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1119 of SEQ ID NO:58, b is an integer of 15 to 1133, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:58, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 49

The translation product of this gene was found to have homology to a 35kd pulmonary surfactant protein, as well as, a GABA-like receptor (See Genbank Accession Nos. P70663, and gil540271, respectively), the latter of which is thought to be important in neuronal function. In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: QKFLCASDGD (SEQ ID NO:534), AEVPLRVRRRHGRPHGPGGRQLALGIPALRSLPGCVPRHHGC SPGYGCLHRRILCLPLILLVYKQRQAASNRRRAQELVRMDSNIQGIENPGF EASPPAQGIPEAKVRHPLSYVAQRQPSSESGRHLLSEPSTPLSPPGPDVFF PSLDPVPDSPNFVIXPXWGTVGCCGWVWGRCI (SEQ ID NO:535), GPGG RQLALGIPALRSLPGCVPRHHGC (SEQ ID NO:536), FEASPPAQGIPEAK VRHPLSYVAQR (SEQ ID NO:537), DMSLGMWQHQQWDMKMDTGPPSQAPD TGHGGETSPPWHALGSPVLPEAALLSDFLVFPQWLWGQACLPTGHRHLPQLPP TSSF SEDLSTG (SEQ ID NO:538), PPSQAPDTGHGGETSPPWHALGS (SEQ ID NO:544), PVDRSSEKLLVGGSWGRWRWPVGRQAWPQSHCGTKRKSDRR AASGKTGEPSACHGGEVSPPCPVSGAWEG GPVSILSH (SEQ ID NO:539), PVDRSSEKLLVGGSWGRWRWPV (SEQ ID NO:540), TKRKSDRRAASG KTGEPSACHGGEV (SEQ ID NO:541). MTSKFGESGTGSRDGKKTSPGPG

GDRGVLGSESRCPDSEGCRRWAT (SEQ ID NO:542), and/or SPGPGGDRGV LGSESRCPD (SEQ ID NO:543). Polynucleotides encoding these polypeptides are also encompassed by the invention.

5 This gene is expressed primarily in hematopoiesis cells such as neutrophils, eosinophils and T cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, blood diseases and/or immune diseases. Similarly, polypeptides and
10 antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the hematopoietic and immune systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., immune, hematopoietic, and cancerous and
15 wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID
20 NO:173 as residues: Ser-44 to Ala-63, Pro-89 to Gly-98, Pro-129 to Trp-137.

The tissue distribution in neutrophils, eosinophils, and T cells indicates polynucleotides and polypeptides corresponding to this gene are useful for treating and diagnosis blood related diseases. Moreover, polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of hematopoietic
25 related disorders such as anemia, pancytopenia, leukopenia, thrombocytopenia or leukemia since stromal cells are important in the production of cells of hematopoietic lineages. The uses include bone marrow cell ex-vivo culture, bone marrow transplantation, bone marrow reconstitution, radiotherapy or chemotherapy of neoplasia. The gene product may also be involved in lymphopoiesis, therefore, it can be
30 used in immune disorders such as infection, inflammation, allergy, immunodeficiency etc. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. The homology to a pulmonary surfactant protein indicates that the protein is useful in enhancing or inhibiting the efficacy of the
35 immune response across mucosal barriers, such as within the gastrointestinal tract, the sinuses, and the lungs. Protein, as well as, antibodies directed against the protein may

show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ
5 ID NO:59 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more
10 polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1476 of SEQ ID NO:59, b is an integer of 15 to 1490, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:59, and where b is greater than or equal to a + 14.

15 FEATURES OF PROTEIN ENCODED BY GENE NO: 50

When tested against U937 cell lines, supernatants removed from cells containing this gene activated the GAS (gamma activating sequence) promoter element. Thus, it is likely that this gene activates promyeloid cells, or more generally, other cells
20 of the immune or central nervous system, through the JAK-STAT signal transduction pathway. GAS is a promoter element found upstream of many genes which are involved in the Jak-STAT pathway. The Jak-STAT pathway is a large, signal transduction pathway involved in the differentiation and proliferation of cells. Therefore, activation of the Jak-STAT pathway, reflected by the binding of the GAS
25 element, can be used to indicate proteins involved in the proliferation and differentiation of cells. The gene encoding the disclosed cDNA is believed to reside on chromosome 1. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 1. In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: HEVQPSYLPNSGLI (SEQ
30 ID NO:545). Polynucleotides encoding these polypeptides are also encompassed by the invention.

This gene is expressed primarily in the central nervous system, adult liver, adult heart, and infant brain.

Therefore, polynucleotides and polypeptides of the invention are useful as
35 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, neural, cardiovascular, or metabolic conditions or disorders. Similarly,

polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the central nervous system, expression of this gene at significantly higher or lower levels may be routinely
5 detected in certain tissues or cell types (e.g., neural, cardiovascular, developmental, metabolic, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, amniotic fluid, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an
10 individual not having the disorder.

The tissue distribution in tissues of the CNS and infant brain, combined with the detected GAS biological activity indicates polynucleotides and polypeptides corresponding to this gene are useful for the treatment, diagnosis, and/or prevention of CNS disorders. Moreover, polynucleotides and polypeptides corresponding to this
15 gene are useful for the detection/treatment of neurodegenerative disease states, behavioral disorders, or inflammatory conditions which include, but are not limited to Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction,
20 aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, depression, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, elevated expression of this gene product in regions of the brain indicates that it plays a role in normal neural function. Potentially, this gene
25 product is involved in synapse formation, neurotransmission, learning, cognition, homeostasis, or neuronal differentiation or survival. The expression within fetal tissue indicates that this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis and treatment of cancer and other proliferative disorders. Similarly, developmental tissues rely on decisions involving cell differentiation and/or
30 apoptosis in pattern formation. Thus this protein may also be involved in apoptosis or tissue differentiation and could again be useful in cancer therapy. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available
35 and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:60 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the

scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1322 of SEQ ID NO:60, b is an integer of 15
5 to 1336, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:60, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 51

10

In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: LRISVLCRETACNWSHHPLDSN (SEQ ID NO:546). Polynucleotides encoding these polypeptides are also encompassed by the invention. The gene encoding the disclosed cDNA is believed to reside on chromosome 18.
15 Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 18.

This gene is expressed in whole brain, embryos, fetal liver and fetal spleen, and melanocytes.

Therefore, polynucleotides and polypeptides of the invention are useful as
20 reagents for identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, neural, immune, hematopoietic, or developmental disorders, particularly mental disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s)
25 or cell type(s). For a number of disorders of the above tissues or cells, particularly of the central nervous system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., neural, immune, hematopoietic, developmental, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, amniotic fluid, urine, synovial fluid and spinal fluid) or another
30 tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:175 as residues: Pro-27 to Lys-42.

35 The tissue distribution in brain indicates polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis, treatment, and/or prevention of mental or neurodegenerative disorders. Alternatively, the expression within fetal

liver/spleen indicates polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of hematopoietic related disorders such as anemia, pancytopenia, leukopenia, thrombocytopenia or leukemia since stromal cells are important in the production of cells of hematopoietic lineages. The uses include

5 bone marrow cell ex-vivo culture, bone marrow transplantation, bone marrow reconstitution, radiotherapy or chemotherapy of neoplasia. The gene product may also be involved in lymphopoiesis, therefore, it can be used in immune disorders such as infection, inflammation, allergy, immunodeficiency etc. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors

10 of various blood lineages, and in the differentiation and/or proliferation of various cell types. The protein may also be useful for the treatment and/or detection of metabolic disorders, which include Tay-Sachs disease, phenylketonuria, galactosemia, hyperlipidemias, porphyrias, and Hurler's syndrome. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy

15 targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:61 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the

20 scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1691 of SEQ ID NO:61, b is an integer of 15 to 1705, where both a and b correspond to the positions of nucleotide residues shown

25 in SEQ ID NO:61, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 52

30 When tested against U937 and fibroblast cell lines, supernatants removed from cells containing this gene activated both the GAS (gamma activating sequence) and EGR1 (early growth response gene 1) promoter elements. Thus, it is likely that this gene activates promyeloid cells, fibroblasts, or more generally, immune or integumentary cells or cell-types, through the JAK-STAT and/or EGR1 signal

35 transduction pathway. GAS is a promoter element found upstream of many genes which are involved in the Jak-STAT pathway. The Jak-STAT pathway is a large, signal transduction pathway involved in the differentiation and proliferation of cells.

Therefore, activation of the Jak-STAT pathway, reflected by the binding of the GAS element, can be used to indicate proteins involved in the proliferation and differentiation of cells. EGR1 is a separate signal transduction pathway from Jak-STAT, genes containing the EGR1 promoter are induced in various tissues and cell types upon
5 activation, leading the cells to undergo differentiation and proliferation. In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: LTVTVRNPGSTHASGRPRRRSGVWARRGLVWQ (SEQ ID NO:547). Polynucleotides encoding these polypeptides are also encompassed by the invention.

This gene is expressed primarily in endometrial stromal cells and fetal brain
10 tissue, and to a lesser extent, in microvascular endothelial cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, reproductive, neural, developmental, or vascular disorders, particularly
15 vascular leak syndrome and inflammation. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the endothelium, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell
20 types (e.g., reproductive, neural, developmental, vascular, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, amniotic fluid, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID
25 NO:176 as residues: Pro-63 to Cys-72, Gly-88 to Cys-93.

The tissue distribution in endometrial stromal cells, infant brain, and microvascular endothelial cells, combined with the detected GAS and EGR1 biological activities, indicates polynucleotides and polypeptides corresponding to this gene are
30 useful for the treatment of various vascular disorders, which include, but are not limited to vascular leak syndrome, microvascular disease, atherosclerosis, aneurysm, stroke, embolism and inflammation. Moreover, the expression within embryonic tissue and other cellular sources marked by proliferating cells indicates that this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis and
35 treatment of cancer and other proliferative disorders. Similarly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation. Thus this protein may also be involved in apoptosis or tissue differentiation and could

again be useful in cancer therapy. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:62 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1017 of SEQ ID NO:62, b is an integer of 15 to 1031, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:62, and where b is greater than or equal to a + 14.

15

FEATURES OF PROTEIN ENCODED BY GENE NO: 53

Contact of cells with supernatant expressing the product of this gene has been shown to increase the permeability of the plasma membrane of HUVEC cells to calcium. Thus, it is likely that the product of this gene is involved in a signal transduction pathway that is initiated when the product binds a receptor on the surface of the plasma membrane of both vascular endothelial cells, in addition to other cell-lines or tissue cell types. Thus, polynucleotides and polypeptides have uses which include, but are not limited to, activating endothelial cells, or more generally, neural or immune cells. Binding of a ligand to a receptor is known to alter intracellular levels of small molecules, such as calcium, potassium and sodium, as well as alter pH and membrane potential. Alterations in small molecule concentration can be measured to identify supernatants which bind to receptors of a particular cell. This protein is homologous to members of the butyrophilin gene family which are thought to play a role in myelin sheath development, in addition to serving as a membrane-specific receptor for cytoplasmic vesicles to the apical plasma membrane. In specific embodiments, polypeptides of the invention comprise the sequence
SAQFSVLGSPGPILAMVGEDADLPCHLFPTMSAETMELKW (SEQ ID NO:574).
Polynucleotides encoding these polypeptides are also encompassed by the invention. In specific embodiments, polypeptides of the invention comprise the sequence
TPCSAQFSVLGSPGPILAMVGEDADLPCHLFPTMSAET (SEQ ID NO:548),
MELKWVSSSLRQVVNVYADGKEVEDRQSAFYRGRTSILRDGITAGKAALRIHN

VTASDSG (SEQ ID NO:549), LEVKG YEDGGIHLECRSTGWYPQPI (SEQ ID NO:550), MASSLAFLLLNFHVSLLL VQLLTPCSAQFSVLGSPGPILAMVGE DADLPCHLFTMSAETMELKWVSSSLRQVVNVYADG (SEQ ID NO:551), RHELSHNRKNGELLIDRLYSVGSDSPMGIPRDIIFTDGFYWNPKVKTLKDRHF

5 WQSIDENGKFPGFPSA QLSCLPPLGPAAHSLSSVFCAWTLWAHPGHGG (SEQ ID NO:552), LLIDRLYSVGSDSPMGIPRDIIFT (SEQ ID NO:553), NPKVKTLKDRHFWQSIDENGKFPGF (SEQ ID NO:554), LGPAAHSLSSVFCAWTLWAHPGH (SEQ ID NO:555). RLQHWVLIFTLEVKG YEDGGIHLECRSTGWYPQPIQWSNAKGENIPAVEAPVVADGVGLYEVAASVIMRGGSGEGVSCIIRNSLL

10 GLEKTASISIADPSSGAPSPGSQPWQGPCLSCCFSPPEVTSCGDNRRK (SEQ ID NO:556), GGIHLECRSTGWYPQPIQWSNAKG (SEQ ID NO:557), PQIQWSNAKGENIPAVEAPVVADGVGL (SEQ ID NO:558), NIPAVEAPVVADGVGLYEVAASVIMRG (SEQ ID NO:559), SGAPSPGSQPWQGPCLSCCFSPPEVT (SEQ ID NO:560), SSSICDHERRLRGGCILHHQKFPPRPGKDSQHFHRRP

15 FFRSAQPWIAALAGTLPILLLLLAGASYFLWRQQKEITALSSEIESEQEMKEMGYAATEREISLRESLQEELKRKKIQYLTRGESSSDTNKSA (SEQ ID NO:561), KDSQHFHRRPFFRSAQPWIAALAGTLP (SEQ ID NO:562), EIESEQEMKEMGYAATEREISLRESLQE (SEQ ID NO:563), VNNMIAFYSDSYVYPHFSGEEMLQMRHLHLVK (SEQ ID NO:564), TPCSAQFSVLGSPGPILAMVGEDADLP

20 CHLFTMSAET (SEQ ID NO:565), KWVSSSLRQVVNVYADGKEVEDR (SEQ ID NO:566), RTSILRDGITAGKAALRIHNVASD (SEQ ID NO:567), CYFQDGDIFYEKALVELKVAALGS (SEQ ID NO:568), GYEDGGIHLECRSTGWYPQPIQ (SEQ ID NO:569), NIPAVEAPVVADGVGLYEVAASV (SEQ ID NO:570), QQKEITALSSEIESEQEMKEM (SEQ ID NO:571), LRESLQEELKRKKIQYLTRGEES (SEQ ID NO:572), and/or GEEMLQMRHLHLVK (SEQ ID NO:573). Polynucleotides encoding these polypeptides are also encompassed by the invention. The gene encoding the disclosed cDNA is believed to reside on chromosome 6. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 6.

30 This gene is expressed primarily in rhabdomyosarcoma, and to a lesser extent, in T cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, muscle, immune, or neural disorders, particularly rhabdomyosarcoma, infectious diseases, or neurodegenerative conditions. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders

of the above tissues or cells, particularly of the immune system expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., muscle, immune, neural, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or
5 cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:177 as residues: Ala-78 to Arg-94.

10 The tissue distribution in rhabdomyosarcoma, the detected calcium flux biological activity, combined with the homology to the butyrophilin gene family indicates polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis or treatment of muscle disorders, which include, but are not limited to, muscular dystrophy, cardiomyopathy, fibroids, myomas, and/or rhabdomyosarcomas.
15 Moreover, the homology to the butyrophilin protein indicates polynucleotides and polypeptides corresponding to this gene are useful for the detection/treatment of neurodegenerative disease states, behavioral disorders, or inflammatory conditions which include, but are not limited to Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating
20 diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, depression, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, elevated
25 expression of this gene product in regions of the brain indicates that it plays a role in normal neural function. Potentially, this gene product is involved in synapse formation, neurotransmission, learning, cognition, homeostasis, or neuronal differentiation or survival. The protein may also show utility in the correction or amelioration of myelin sheath deficiencies in developing and mature neurons and neural-cell types. Protein, as
30 well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:63 and may have been publicly available prior to conception of the present
35 invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more

polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1575 of SEQ ID NO:63, b is an integer of 15 to 1589, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:63, and where b is greater than or equal to a + 14.

5

FEATURES OF PROTEIN ENCODED BY GENE NO: 54

10 In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: PQGGLTLPSVWG (SEQ ID NO:575), GGPCHLWLLGPRRT QLPGRRASLPFRSQGELTQAFLLGLWKHQMPALTQEQQVRAERRREAVRMEI PGLFFASLANWGLLYRTSQDFISPYLCAAPSTPHPLGGP (SEQ ID NO:576), GPRRTQLPGRRASLPFRSQGELT (SEQ ID NO:577), QMPALTQEQQVRAER RREAVRMEI (SEQ ID NO:578), and/or ANWGLLYRTSQDFISPYLCAAPSTP (SEQ
15 ID NO:579). Polynucleotides encoding these polypeptides are also encompassed by the invention. The gene encoding the disclosed cDNA is believed to reside on chromosome 5. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 5.

20 This gene is expressed primarily in brain, and to a lesser extent, in testes tumor. Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, neural, endocrine, or reproductive disorders, particularly depression and infertility disorders. Similarly, polypeptides and antibodies directed to these
25 polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the endocrine and nervous systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., neural, endocrine, reproductive, and cancerous and wounded tissues) or
30 bodily fluids (e.g., serum, plasma, urine, seminal fluid, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

35 Preferred epitopes include those comprising a sequence shown in SEQ ID NO:178 as residues: Thr-26 to Glu-33.

The tissue distribution in brain indicates polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of depression and

other endocrine-related disorders. Moreover, polynucleotides and polypeptides corresponding to this gene are useful for the detection/treatment of neurodegenerative disease states, behavioral disorders, or inflammatory conditions which include, but are not limited to Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, 5 Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, depression, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in 10 feeding, sleep patterns, balance, and perception. In addition, elevated expression of this gene product in regions of the brain indicates that it plays a role in normal neural function. Potentially, this gene product is involved in synapse formation, neurotransmission, learning, cognition, homeostasis, or neuronal differentiation or survival. The protein product may also be useful in the treatment, detection, and/or 15 prevention of a variety of reproductive disorders which include, but are not limited to, the treatment of male infertility, and/or could be used as a male contraceptive. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available 20 and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:64 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more 25 polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1074 of SEQ ID NO:64, b is an integer of 15 to 1088, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:64, and where b is greater than or equal to a + 14.

30

FEATURES OF PROTEIN ENCODED BY GENE NO: 55

The translation product of this gene was found to have homology to the conserved R166.2 protein from *Caenorhabditis elegans* (See Genbank Accession 35 No.gil949849), which is thought to play an important role in the regulation of cellular function and processes. In specific embodiments, polypeptides of the invention comprise the sequence: LSFKDKSTYIESSTKVYDDMAFRYLSWILFPLL (SEQ ID

NO:580), CYAVYSLLYLEHKGWYSWVLSM (SEQ ID NO:590), LLTFGFITMTPQ
 LFINYKLKSV AHPWRMLT (SEQ ID NO:581), TYKALNTFIDDLFAFVIKMP
 VMYRIGCLRD (SEQ ID NO:582), DVVFFIYLYQRWIYRVDPTRVNEFGMSGED
 (SEQ ID NO:583), VAGIFPRLSFKDKSTYIESSTKVYDDMAFRYLSWILFPLLG
 5 CYA (SEQ ID NO:584), PWVAGIFPRLSFKDKSTYIESSTKVYDD (SEQ ID
 NO:586), AGEDSCHPVLSVQPDVHDLGWQESSPA YPSRTSPRISSPRPKC
 MMIWHSGTCPGSSSR SWAAMPSTVFCTWSTRAGTPGCSACSTASC (SEQ ID
 NO:587), LSVQPDVHDLGWQESSPA YPSRTSPRISSP (SEQ ID NO:588), GSSSR
 SWAAMPSTVFCTWSTRAGTP (SEQ ID NO:589), and/or WAAMPSTVFCTWS
 10 TRAGTP (SEQ ID NO:585). Polynucleotides encoding these polypeptides are also
 encompassed by the invention. The gene encoding the disclosed cDNA is believed to
 reside on chromosome 19. Accordingly, polynucleotides related to this invention are
 useful as a marker in linkage analysis for chromosome 19.

This gene is expressed primarily in colon, smooth muscle and fetal bone.

15 Therefore, polynucleotides and polypeptides of the invention are useful as
 reagents for differential identification of the tissue(s) or cell type(s) present in a
 biological sample and for diagnosis of diseases and conditions which include, but are
 not limited to, gastrointestinal or vascular disorders, and abnormal muscular-skeletal
 development, including proliferative conditions such as cancer. Similarly, polypeptides
 20 and antibodies directed to these polypeptides are useful in providing immunological
 probes for differential identification of the tissue(s) or cell type(s). For a number of
 disorders of the above tissues or cells, particularly of the immune and muscular-skeletal
 system, expression of this gene at significantly higher or lower levels may be routinely
 detected in certain tissues or cell types (e.g., gastrointestinal, muscle, skeletal, vascular,
 25 developmental, and cancerous and wounded tissues) or bodily fluids (e.g., serum,
 plasma, amniotic fluid, urine, synovial fluid and spinal fluid) or another tissue or cell
 sample taken from an individual having such a disorder, relative to the standard gene
 expression level, i.e., the expression level in healthy tissue or bodily fluid from an
 individual not having the disorder.

30 Preferred epitopes include those comprising a sequence shown in SEQ ID
 NO:179 as residues: Ser-128 to Thr-133, Thr-158 to Thr-166, Leu-168 to Gly-175,
 Ala-179 to Asp-196.

The tissue distribution in colon and fetal bone indicates polynucleotides and
 polypeptides corresponding to this gene are useful for the treatment and diagnosis of
 35 abnormal bone formation, and/or various proliferative conditions (e.g. tumors),
 particularly of the gastrointestinal system. Moreover, the expression within smooth
 muscle tissue indicates polynucleotides and polypeptides corresponding to this gene are

useful for the detection, treatment, and/or prevention of a variety of vascular disorders, which include, but are not limited to the following: embolism, atherosclerosis, microvascular disease, aneurysm, stroke, and vascular leak syndrome. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:65 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1242 of SEQ ID NO:65, b is an integer of 15 to 1256, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:65, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 56

In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: LGEFLSSQCFLP (SEQ ID NO:591). Polynucleotides encoding these polypeptides are also encompassed by the invention.

This gene is expressed primarily in brain frontal cortex.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, neural disorders, particularly neurological or neurodegenerative disorders and diseases. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the brain, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., neural, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:180 as residues: Ala-122 to Gly-128.

5 The tissue distribution in brain frontal cortex indicates polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of some neurological diseases such as depression. Moreover, polynucleotides and polypeptides corresponding to this gene are useful for the detection/treatment of neurodegenerative disease states, behavioral disorders, or inflammatory conditions which include, but are not limited to Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating
10 diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, depression, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, elevated
15 expression of this gene product in regions of the brain indicates that it plays a role in normal neural function. Potentially, this gene product is involved in synapse formation, neurotransmission, learning, cognition, homeostasis, or neuronal differentiation or survival. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

20 Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:66 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome.
25 Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1588 of SEQ ID NO:66, b is an integer of 15 to 1602, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:66, and where b is greater than or equal to a + 14.

30

FEATURES OF PROTEIN ENCODED BY GENE NO: 57

35 When tested against U937 cell lines, supernatants removed from cells containing this gene activated the GAS (gamma activating sequence) promoter element. Thus, it is likely that this gene activates pro-myeloid cells, or more generally, immune or hematopoietic cells, through the JAK-STAT signal transduction pathway. GAS is a

promoter element found upstream of many genes which are involved in the Jak-STAT pathway. The Jak-STAT pathway is a large, signal transduction pathway involved in the differentiation and proliferation of cells. Therefore, activation of the Jak-STAT pathway, reflected by the binding of the GAS element, can be used to indicate proteins involved in the proliferation and differentiation of cells. In specific embodiments, polypeptides of the invention comprise the following amino acid sequence:

5 RSRNRVAMGMWASLDALWE (SEQ ID NO:592), PRVRCQQRAEGGMGAG
IGVGPSERTDIAVTPRGRSEGASVGVAPVHAEGAGGTGWPWGCGHRWTLCG
RCR PRSVSSGPCCSFPGQCIFGRPS (SEQ ID NO:593), GGMGAGIGVGPSER
10 TDIIVTPRGR (SEQ ID NO:594), GCGHRWTLCGRRCR PRSVSSGPCCSFP (SEQ
ID NO:595), and/or KKHGF NQQTGLGFFTWKYNKNKNLV (SEQ ID NO:596).
Polynucleotides encoding these polypeptides are also encompassed by the invention.
The gene encoding the disclosed cDNA is believed to reside on chromosome 1.
Accordingly, polynucleotides related to this invention are useful as a marker in linkage
15 analysis for chromosome 1.

This gene is expressed primarily in synovial cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, skeletal afflictions, particularly rheumatoid arthritis or autoimmune conditions. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., skeletal, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

30 Preferred epitopes include those comprising a sequence shown in SEQ ID NO:181 as residues: Gln-27 to Val-39, Glu-50 to Arg-56.

The restricted tissue distribution in synovium, combined with the detected GAS biological activity, indicates polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of rheumatoid arthritis since synovial fibroblasts are associated with the synovium and cartilage. Moreover, polynucleotides and polypeptides corresponding to this gene are useful in the detection and treatment of disorders and conditions afflicting the skeletal system, in particular osteoporosis, bone

35

cancer, as well as, disorders afflicting connective tissues (e.g. arthritis, trauma, tendonitis, chondromalacia and inflammation), such as in the diagnosis or treatment of various autoimmune disorders such as rheumatoid arthritis, lupus, scleroderma, and dermatomyositis as well as dwarfism, spinal deformation, and specific joint abnormalities as well as chondrodysplasias (i.e. spondyloepiphyseal dysplasia congenita, familial osteoarthritis, Atelosteogenesis type II, metaphyseal chondrodysplasia type Schmid). The protein may also be useful in the modulation of the immune response to regions of inflammation, or in inhibiting or ameliorating autoimmune responses. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:67 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 924 of SEQ ID NO:67, b is an integer of 15 to 938, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:67, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 58

When tested against U937 cell lines, supernatants removed from cells containing this gene activated the GAS (gamma activating sequence) promoter element. Thus, it is likely that this gene activates pro-myeloid cells, or more generally immune or hematopoietic cells, in addition to other cells or cell-types, through the JAK-STAT signal transduction pathway. GAS is a promoter element found upstream of many genes which are involved in the Jak-STAT pathway. The Jak-STAT pathway is a large, signal transduction pathway involved in the differentiation and proliferation of cells. Therefore, activation of the Jak-STAT pathway, reflected by the binding of the GAS element, can be used to indicate proteins involved in the proliferation and differentiation of cells. In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: PKLLPCSPAEGHTSLGPLLPF (SEQ ID NO:597), ASLELXPS KSQLSTEWGFTWIVGLGMSPSTALWTECTCTPFLVLLSHASGHFFWLSPLAS

LVIPPVTDRK (SEQ ID NO:598), WGFTWIVGLGMSPSTALWTECTCTPFLVL
LSH (SEQ ID NO:599), VAVGVCREDMVGITDRSKMSPDVGIWAIYWSAAGY
WPLIGFPGTPTQQEPALHRVGVYLDRTGNVSFYSAVDGVHLHTFSCS
SVSRLRPFFLVESISIFSHSTSD (SEQ ID NO:600), ITDRSKMSPDVGIWAIYW
5 SAAGYWPLI (SEQ ID NO:601), and/or RGTGNVSFYSAVDGVHLHTFSCSSV
SRLRP (SEQ ID NO:602). Polynucleotides encoding these polypeptides are also
encompassed by the invention. The gene encoding the disclosed cDNA is believed to
reside on chromosome 7. Accordingly, polynucleotides related to this invention are
useful as a marker in linkage analysis for chromosome 7.

10 This gene is expressed primarily in fetal tissues, and to a lesser extent, in liver
cells.

Therefore, polynucleotides and polypeptides of the invention are useful as
reagents for differential identification of the tissue(s) or cell type(s) present in a
biological sample and for diagnosis of diseases and conditions which include, but are
15 not limited to, developmental or liver diseases, such as hepatocellular carcinoma.
Similarly, polypeptides and antibodies directed to these polypeptides are useful in
providing immunological probes for differential identification of the tissue(s) or cell
type(s). For a number of disorders of the above tissues or cells, particularly of the
immune and hepatic systems, expression of this gene at significantly higher or lower
20 levels may be routinely detected in certain tissues or cell types (e.g., developmental,
hepatic, metabolic, and cancerous and wounded tissues) or bodily fluids (e.g., serum,
plasma, bile, breast milk, urine, synovial fluid and spinal fluid) or another tissue or cell
sample taken from an individual having such a disorder, relative to the standard gene
expression level, i.e., the expression level in healthy tissue or bodily fluid from an
25 individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID
NO:182 as residues: Pro-30 to Gln-37, Arg-39 to Ser-45, Arg-74 to Arg-85.

The tissue distribution in liver, combined with the detected GAS biological
activity indicates polynucleotides and polypeptides corresponding to this gene are useful
30 for the treatment or diagnosis of hepatic conditions such as hepatocellular carcinoma.
Moreover, the expression within embryonic tissue and other cellular sources marked by
proliferating cells indicates that this protein may play a role in the regulation of cellular
division, and may show utility in the diagnosis and treatment of cancer and other
proliferative disorders. Similarly, developmental tissues rely on decisions involving cell
35 differentiation and/or apoptosis in pattern formation. Thus this protein may also be
involved in apoptosis or tissue differentiation and could again be useful in cancer

therapy. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:68 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1571 of SEQ ID NO:68, b is an integer of 15 to 1585, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:68, and where b is greater than or equal to a + 14.

15 FEATURES OF PROTEIN ENCODED BY GENE NO: 59

In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: GTRGLQNRTE (SEQ ID NO:603). Polynucleotides encoding these polypeptides are also encompassed by the invention.

20 This gene is expressed primarily in prostate, and to a lesser extent, in tonsil and fetal lung.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, reproductive, immune, developmental, or pulmonary disorders and/or diseases. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., reproductive, immune, developmental, pulmonary, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, amniotic fluid, pulmonary surfactant or sputum, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:183 as residues: Lys-32 to Lys-38.

- The tissue distribution in prostate indicates polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of cancers, particularly of the prostate. The expression within tonsils indicates polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. The expression also indicates a role in regulating the proliferation; survival; differentiation; and/or activation of hematopoietic cell lineages, including blood stem cells. Moreover, the expression within fetal tissue indicates that this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis and treatment of cancer and other proliferative disorders.
- Similarly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation. Thus this protein may also be involved in apoptosis or tissue differentiation and could again be useful in cancer therapy. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.
- Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:69 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome.
- Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1662 of SEQ ID NO:69, b is an integer of 15 to 1676, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:69, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 60

- In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: ELSGLG (SEQ ID NO:604). Polynucleotides encoding these polypeptides are also encompassed by the invention.

This gene is expressed primarily in the brain.

- Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, central nervous system disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for

differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the central nervous system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., neural, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:184 as residues: Tyr-15 to Lys-21, Pro-62 to Phe-68.

The tissue distribution in brain indicates polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis, treatment, and/or prevention of CNS disorders (such as Parkinson's disease). Moreover, polynucleotides and polypeptides corresponding to this gene are useful for the detection/treatment of neurodegenerative disease states, behavioral disorders, or inflammatory conditions which include, but are not limited to Alzheimer's Disease, Huntington's Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, depression, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, elevated expression of this gene product in regions of the brain indicates that it plays a role in normal neural function. Potentially, this gene product is involved in synapse formation, neurotransmission, learning, cognition, homeostasis, or neuronal differentiation or survival. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:70 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1330 of SEQ ID NO:70, b is an integer of 15 to 1344, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:70, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 61

5 The gene encoding the disclosed cDNA is believed to reside on chromosome 3. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 3.

 This gene is expressed primarily in the brain.

 Therefore, polynucleotides and polypeptides of the invention are useful as
10 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, CNS diseases, such as Alzheimers and Parkinson's disease. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a
15 number of disorders of the above tissues or cells, particularly of the central nervous system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., neural, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to
20 the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

 Preferred epitopes include those comprising a sequence shown in SEQ ID NO:185 as residues: Asp-44 to Cys-53, Asp-56 to Lys-66, Ser-78 to Lys-84.

 The tissue distribution in brain tissue indicates polynucleotides and polypeptides
25 corresponding to this gene are useful for the diagnosis, treatment, and/or prevention of CNS diseases such as Alzheimers and Parkinson's disease. Moreover, polynucleotides and polypeptides corresponding to this gene are useful for the detection/treatment of neurodegenerative disease states, behavioral disorders, or inflammatory conditions which include, but are not limited to Huntington's Disease, Tourette Syndrome,
30 meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, depression, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance,
35 and perception. In addition, elevated expression of this gene product in regions of the brain indicates that it plays a role in normal neural function. Potentially, this gene product is involved in synapse formation, neurotransmission, learning, cognition,

homeostasis, or neuronal differentiation or survival. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:71 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1460 of SEQ ID NO:71, b is an integer of 15 to 1474, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:71, and where b is greater than or equal to a + 14.

15

FEATURES OF PROTEIN ENCODED BY GENE NO: 62

In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: MDDIKI (SEQ ID NO:605), NFCVSKNTFNRVKRPIKWVKIF ANDISCKRLISRIHKEILPFNNKKQPDFKVKKSRK (SEQ ID NO:606), FNRVKR PIKWVKIFANDISCKRLISRIHKE (SEQ ID NO:607), ETQMANKYMKRCSTL (SEQ ID NO:608), VIRELQVKATRCHYTPIKWSKSKTLISSNADEYVEPTRTLI HCWWKCKIVQPLCKTAW (SEQ ID NO:609), and/or ATRCHYTPIKWSKSKT LISSN (SEQ ID NO:610). Polynucleotides encoding these polypeptides are also encompassed by the invention.

This gene is expressed primarily in duodenum, and to a lesser extent, in brain frontal cortex.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, gastrointestinal, neural, or endocrine disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the endocrine system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., gastrointestinal, neural, endocrine, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid and spinal

fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in colon indicates polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of some gastrointestinal disorders, particularly cancers. Moreover, polynucleotides and polypeptides corresponding to this gene are useful for the detection/treatment of neurodegenerative disease states, behavioral disorders, or inflammatory conditions which include, but are not limited to Alzheimer's Disease, Parkinson's Disease, 10 Huntington's Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, depression, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including 15 disorders in feeding, sleep patterns, balance, and perception. In addition, elevated expression of this gene product in regions of the brain indicates that it plays a role in normal neural function. Potentially, this gene product is involved in synapse formation, neurotransmission, learning, cognition, homeostasis, or neuronal differentiation or survival. Protein, as well as, antibodies directed against the protein may show utility as 20 a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:72 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the 25 scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1998 of SEQ ID NO:72, b is an integer of 15 to 2012, where both a and b correspond to the positions of nucleotide residues shown 30 in SEQ ID NO:72, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 63

35 This gene is expressed primarily in bone marrow.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a

biological sample and for diagnosis of diseases and conditions which include, but are not limited to, leukemia, immune deficiency syndromes, and other immune related diseases. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., immune, cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of diseases of bone marrow, such as leukemia, bone cancer and immune deficiency syndrome. Furthermore, the polypeptides or polynucleotides are also useful to enhance or protect proliferation, differentiation, and functional activation of hematopoietic progenitor cells (e.g., bone marrow cells), useful in treating cancer patients undergoing chemotherapy or patients undergoing bone marrow transplantation. The uses include bone marrow cell ex-vivo culture, bone marrow transplantation, bone marrow reconstitution, radiotherapy or chemotherapy of neoplasia. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:73 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1253 of SEQ ID NO:73, b is an integer of 15 to 1267, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:73, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 64

When tested against sensory neuron cell lines, supernatants removed from cells containing this gene activated the EGR1 assay. Thus, it is likely that this gene activates
5 sensory neuronal cells through a signal transduction pathway. Early growth response 1 (EGR1) is a promoter associated with certain genes that induces various tissues and cell types upon activation, leading the cells to undergo differentiation and proliferation.

This gene is expressed primarily in the testes.

Therefore, polynucleotides and polypeptides of the invention are useful as
10 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, reproductive system-related diseases. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders
15 of the above tissues or cells, particularly of the reproductive system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., testes, cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene
20 expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in testes indicates polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis, treatment, and/or prevention of reproductive system-related diseases. Furthermore, the tissue distribution indicates that
25 polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of conditions concerning proper testicular function (e.g. endocrine function, sperm maturation), as well as cancer. Therefore, this gene product is useful in the treatment of male infertility and/or impotence. This gene product is also useful in assays designed to identify binding agents, as such agents (antagonists) are useful as
30 male contraceptive agents. Similarly, the protein is believed to be useful in the treatment and/or diagnosis of testicular cancer. The testes are also a site of active gene expression of transcripts that may be expressed, particularly at low levels, in other tissues of the body. Therefore, this gene product may be expressed in other specific tissues or organs where it may play related functional roles in other processes, such as hematopoiesis,
35 inflammation, bone formation, and kidney function, to name a few possible target indications.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:74 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1734 of SEQ ID NO:74, b is an integer of 15 to 1748, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:74, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 65

This gene is expressed primarily in synovial fibroblasts.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, rheumatoid arthritis. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., immune, cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The restricted tissue distribution indicates polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of rheumatoid arthritis, since synovial fibroblasts are associated with the synovium and cartilage. In addition, the expression of this gene product in synovium indicates a role in the detection and treatment of disorders and conditions afflicting the skeletal system, in particular osteoporosis as well as disorders afflicting connective tissues (e.g. arthritis, trauma, tendonitis, chondromalacia and inflammation), such as in the diagnosis or treatment of various autoimmune disorders such as rheumatoid arthritis, lupus, scleroderma, and dermatomyositis as well as dwarfism, spinal deformation, and

specific joint abnormalities as well as chondrodysplasias (ie. spondyloepiphyseal dysplasia congenita, familial osteoarthritis, Atelosteogenesis type II, metaphyseal chondrodysplasia type Schmid). Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above
5 listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:75 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the
10 scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1556 of SEQ ID NO:75, b is an integer of 15 to 1570, where both a and b correspond to the positions of nucleotide residues shown
15 in SEQ ID NO:75, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 66

20 This gene is expressed primarily in ovarian cancer, and to a lesser extent in fetal tissues such as fetal liver and fetal brain.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are
25 not limited to, cancers, particularly of the ovary. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell
30 types (e.g., ovary, fetal, cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

35 Preferred epitopes include those comprising a sequence shown in SEQ ID NO:190 as residues: Pro-28 to Gln-33.

The tissue distribution in ovarian cancer tissue indicates polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of cancers; e.g., ovarian cancer, as well as other tissues where expression has been indicated. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:76 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 510 of SEQ ID NO:76, b is an integer of 15 to 524, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:76, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 67

This gene is expressed primarily in testes.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, male reproductive or hormonal related disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the male reproductive system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., testes, cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:191 as residues: Pro-68 to Asp-73, Gln-92 to Glu-107, Gln-120 to Lys-126.

The tissue distribution in testes indicates polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of male

reproductive or hormonal disorders. Furthermore, the tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of conditions concerning proper testicular function (e.g. endocrine function, sperm maturation), as well as cancer. Therefore, this gene product is useful in the treatment of male infertility and/or impotence. This gene product is also useful in assays designed to identify binding agents, as such agents (antagonists) are useful as male contraceptive agents. Similarly, the protein is believed to be useful in the treatment and/or diagnosis of testicular cancer. The testes are also a site of active gene expression of transcripts that may be expressed, particularly at low levels, in other tissues of the body. Therefore, this gene product may be expressed in other specific tissues or organs where it may play related functional roles in other processes, such as hematopoiesis, inflammation, bone formation, and kidney function, to name a few possible target indications. Protein, as well as, antibodies directed against the protein may show utility as a tissue-specific marker and/or immunotherapy target for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:77 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1292 of SEQ ID NO:77, b is an integer of 15 to 1306, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:77, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 68

When tested against U937 Myeloid cell lines, supernatants removed from cells containing this gene activated the GAS assay. Thus, it is likely that this gene activates myeloid cells through the Jak-STAT signal transduction pathway. The gamma activating sequence (GAS) is a promoter element found upstream of many genes which are involved in the Jak-STAT pathway. The Jak-STAT pathway is a large, signal transduction pathway involved in the differentiation and proliferation of cells. Therefore, activation of the Jak-STAT pathway, reflected by the binding of the GAS element, can be used to indicate proteins involved in the proliferation and differentiation of cells.

This gene is expressed primarily in human tonsils.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., tonsils, cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of immune disorders. Expression of this gene product in tonsils indicates a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the gene or protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:78 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more

polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1465 of SEQ ID NO:78, b is an integer of 15 to 1479, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:78, and where b is greater than or equal to a + 14.

5

FEATURES OF PROTEIN ENCODED BY GENE NO: 69

10 This gene is expressed primarily in human thymus and six week old human embryo.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, endocrine diseases and leukemia. Similarly, polypeptides and antibodies
15 directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the endocrine and immune systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., endocrine, immune, cancerous and wounded
20 tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in thymus and developing embryonic tissues indicates
25 that polynucleotides and polypeptides corresponding to this gene are useful for the treatment of leukemia or other immune diseases, especially those which are involved in fetal development. Furthermore, the tissue distribution in thymus and developing embryonic tissues indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection, treatment, and/or prevention of various endocrine
30 disorders and cancers, particularly Addison's disease, Cushing's Syndrome, and disorders and/or cancers of the pancreas (e.g. diabetes mellitus), adrenal cortex, ovaries, pituitary (e.g., hyper-, hypopituitarism), thyroid (e.g. hyper-, hypothyroidism), parathyroid (e.g. hyper-, hypoparathyroidism), hypothalamus, and testes. Protein, as well as, antibodies directed against the protein may show utility as a
35 tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ

ID NO:79 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome.

Accordingly, preferably excluded from the present invention are one or more
5 polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1780 of SEQ ID NO:79, b is an integer of 15 to 1794, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:79, and where b is greater than or equal to a + 14.

10

FEATURES OF PROTEIN ENCODED BY GENE NO: 70

This gene is expressed primarily in adult pulmonary tissue.

Therefore, polynucleotides and polypeptides of the invention are useful as
15 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, diseases of the cardiopulmonary system including asthma, bronchitis, apnea, enlarged heart, arrhythmia, strokes and heart attacks. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes
20 for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the cardiopulmonary system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., pulmonary, cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell
25 sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:194 as residues: Pro-27 to Leu-41.

30 The tissue distribution in pulmonary tissues indicates polynucleotides and polypeptides corresponding to this gene are useful for the treatment or diagnosis of diseases such as arrhythmia, apnea, asthma and possibly for the early detection and prevention of patients likely to have strokes or heart attacks. Furthermore, the tissue distribution in pulmonary tissues indicates polynucleotides and polypeptides
35 corresponding to this gene are useful for the detection and treatment of disorders associated with developing lungs, particularly in premature infants where the lungs are the last tissues to develop. Additionally, the tissue distribution indicates polynucleotides

and polypeptides corresponding to this gene are useful for the diagnosis and intervention of lung tumors. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and immunotherapy targets for the above listed tumors and tissues.

- 5 Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:80 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome.
- 10 Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1266 of SEQ ID NO:80, b is an integer of 15 to 1280, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:80, and where b is greater than or equal to a + 14.

15

FEATURES OF PROTEIN ENCODED BY GENE NO: 71

- 20 When tested against K562 leukemia cell lines, supernatants removed from cells containing this gene activated the ISRE assay. Thus, it is likely that this gene activates leukemia cells through the Jak-STAT signal transduction pathway. The interferon-sensitive response element is a promoter element found upstream of many genes which are involved in the Jak-STAT pathway. The Jak-STAT pathway is a large, signal transduction pathway involved in the differentiation and proliferation of cells.
- 25 Therefore, activation of the Jak-STAT pathway, reflected by the binding of the ISRE element, can be used to indicate proteins involved in the proliferation and differentiation of cells. Furthermore, contact of cells with supernatant expressing the product of this gene increases the permeability of THP-1 monocyte cells to calcium. Thus, it is likely that the product of this gene is involved in a signal transduction pathway that is initiated
- 30 when the product of this gene binds a receptor on the surface of the monocyte cell. Thus, polynucleotides and polypeptides have uses which include, but are not limited to, activating monocyte cells.

This gene is expressed primarily in adult human spleen and adult human testis.

- 35 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune disorders. Similarly, polypeptides and antibodies directed to

these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., spleen, testes, cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

10 Preferred epitopes include those comprising a sequence shown in SEQ ID NO:195 as residues: Pro-32 to Gly-39.

The tissue distribution in spleen, in addition to the biological activity data, indicates polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of immune disorders. Furthermore, this gene may play a role in the survival, proliferation, and/or differentiation of hematopoietic cells in general, and may be of use in the augmentation of the numbers of stem cells and committed progenitors. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:81 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 960 of SEQ ID NO:81, b is an integer of 15 to 974, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:81, and where b is greater than or equal to a + 14.

30

FEATURES OF PROTEIN ENCODED BY GENE NO: 72

In specific embodiments, polypeptides of the invention comprise the following amino acid sequence:

35 ELSGLVIITAWIILCHSSSKNPVGGRIQLAIAIVITLFPFISWVYIYINKEMRSSWP
THCKTVI (SEQ ID NO:611), QCPQGTETEAGVSVPPRKEGGGPYVAGLTAPHVA
GLTAPRRVLRAMAPALWRACNGL (SEQ ID NO:612), HSSSKNPVGGRIQLA

IAIVITLFPFISWVYTY (SEQ ID NO:613), and/or RKEGGGPYVAGLTAPHVA
GLTAPRRVLRAMAP (SEQ ID NO:614). Polynucleotides encoding these
polypeptides are also encompassed by the invention.

5 This gene is expressed primarily in liver tissues, and to a lesser extent in t-cell
lymphoma.

Therefore, polynucleotides and polypeptides of the invention are useful as
reagents for differential identification of the tissue(s) or cell type(s) present in a
biological sample and for diagnosis of diseases and conditions which include, but are
not limited to, hepatitis, sclerosis of the liver and cancer of the liver. Similarly,
10 polypeptides and antibodies directed to these polypeptides are useful in providing
immunological probes for differential identification of the tissue(s) or cell type(s). For a
number of disorders of the above tissues or cells, particularly of the immune system,
expression of this gene at significantly higher or lower levels may be routinely detected
in certain tissues or cell types (e.g., immune, cancerous and wounded tissues) or bodily
15 fluids (e.g., serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or
cell sample taken from an individual having such a disorder, relative to the standard
gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an
individual not having the disorder.

The tissue distribution in liver indicates polynucleotides and polypeptides
20 corresponding to this gene are useful for the diagnosis and possible treatment of
diseases of the liver. Since it is primarily found in the liver, and with the additional
expression seen in T-cells, it most likely deals with the immune response in the liver,
for example to diseases like hepatitis, sclerosis and hepatocellular carcinoma. More
generally, the tissue distribution in liver indicates polynucleotides and polypeptides
25 corresponding to this gene are useful for the detection and treatment of liver disorders
and cancers (e.g. hepatoblastoma, jaundice, hepatitis, liver metabolic diseases and
conditions that are attributable to the differentiation of hepatocyte progenitor cells).
Protein, as well as, antibodies directed against the protein may show utility as a tumor
marker and/or immunotherapy targets for the above listed tissues.

30 Many polynucleotide sequences, such as EST sequences, are publicly available
and accessible through sequence databases. Some of these sequences are related to SEQ
ID NO:82 and may have been publicly available prior to conception of the present
invention. Preferably, such related polynucleotides are specifically excluded from the
scope of the present invention. To list every related sequence is cumbersome.
35 Accordingly, preferably excluded from the present invention are one or more
polynucleotides comprising a nucleotide sequence described by the general formula of
a-b, where a is any integer between 1 to 1941 of SEQ ID NO:82, b is an integer of 15

to 1955, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:82, and where b is greater than or equal to a + 14.

5 FEATURES OF PROTEIN ENCODED BY GENE NO: 73

The gene encoding the disclosed cDNA is thought to reside on chromosome 10. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 10.

- 10 This gene is expressed primarily in myeloid progenitor cells, and to a lesser extent in leukemic cells and eosinophils.

- Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, leukemia and other blood diseases. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the hematopoiesis and immune systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., immune, cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

- 25 The tissue distribution in immune tissues indicates polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of leukemia. Furthermore, the polypeptides or polynucleotides are also useful to enhance or protect proliferation, differentiation, and functional activation of hematopoietic progenitor cells (e.g., bone marrow cells), useful in treating cancer patients undergoing chemotherapy or patients undergoing bone marrow transplantation. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

- 30 Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:83 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome.

Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 624 of SEQ ID NO:83, b is an integer of 15 to 638, where both a and b correspond to the positions of nucleotide residues shown in
5 SEQ ID NO:83, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 74

10 The translation product of this gene shares sequence homology with "neurogenic secreted signaling protein (brn)" (see gil1150971) from Drosophila melanogaster which is thought to be important in the normal development of brain tissue. In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: PGRPTRPAXAGLSSGGAAQEAPQADPRPWLAR (SEQ ID
15 NO:615). Polynucleotides encoding these polypeptides are also encompassed by the invention.

This gene is expressed primarily in the placenta and early embryonic tissue. Northern data has demonstrated that this gene is expressed in brain, stomach and colorectal adenocarcinoma.

20 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, several types of disorders of the brain including epilepsy, mood disorders, any of a variety of types of mental retardation, and addictive disorders
25 including alcoholism. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the central nervous system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., brain,
30 stomach, colon, placental, embryonic, cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

35 Preferred epitopes include those comprising a sequence shown in SEQ ID NO:198 as residues: Gln-37 to Ala-42, Thr-51 to Ala-57, Pro-71 to His-79, Glu-124 to Arg-137, Ser-151 to Val-159.

The tissue distribution and homology to *Drosophila melanogaster* putative neurogenic secreted signaling protein (bn) indicates polynucleotides and polypeptides corresponding to this gene are useful for the treatment of various brain disorders as well as pre-natal testing for neuropathological conditions, such as Alzheimers Disease, Parkinsons Disease, Huntingtons Disease, Tourette Syndrome, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, the gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the developing embryo, or sexually-linked disorders. Furthermore, the tissue distribution indicates polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and/or treatment of disorders of the placenta. Specific expression within the placenta indicates that this gene product may play a role in the proper establishment and maintenance of placental function. Alternately, this gene product may be produced by the placenta and then transported to the embryo, where it may play a crucial role in the development and/or survival of the developing embryo or fetus. Expression of this gene product in a vascular-rich tissue such as the placenta also indicates that this gene product may be produced more generally in endothelial cells or within the circulation. In such instances, it may play more generalized roles in vascular function, such as in angiogenesis. It may also be produced in the vasculature and have effects on other cells within the circulation, such as hematopoietic cells. It may serve to promote the proliferation, survival, activation, and/or differentiation of hematopoietic cells, as well as other cells throughout the body. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:84 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 845 of SEQ ID NO:84, b is an integer of 15 to 859, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:84, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 75

5 The translation product of this gene shares sequence homology with a fat-specific secreted protein.

This gene is expressed primarily in the epididymus.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, metabolic disorders and male infertility. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the metabolic and reproductive systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., epididymus, cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

20 Preferred epitopes include those comprising a sequence shown in SEQ ID NO:199 as residues: Tyr-21 to Asp-40, Ser-58 to Arg-64, Thr-71 to Ser-76, Ser-106 to Thr-112.

Homology to a fat-specific gene indicates that this gene may also play a role in the treatment and/or detection of metabolic disorders such as obesity, diabetes, anorexia nervosa and bulimia. In addition, its expression primarily in the epididymus indicates a role in the treatment/detection of male fertility disorders such as infertility, low sperm count, spermatorrhea and spermiation. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

30 Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:85 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome.

35 Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1115 of SEQ ID NO:85, b is an integer of 15

to 1129, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:85, and where b is greater than or equal to a + 14.

5 FEATURES OF PROTEIN ENCODED BY GENE NO: 76

The translation product of this gene shares sequence homology with Slit, a secreted Drosophila protein which plays a role in the development of axon pathway development in the central nervous system. The Slit protein is necessary for the normal
10 development of the midline of the CNS, particularly the midline glial cells, and for the concomitant formation of the commissural axon pathways. The process is dependent on the level of SLIT protein expression. It appears that the SLIT protein is excreted by the midline glial cells, where it is synthesised and is eventually associated with the surfaces of axons that traverse them. Contact of cells with supernatant expressing the
15 product of this gene increases the permeability of THP-1 monocyte cells to calcium. Thus, it is likely that the product of this gene is involved in a signal transduction pathway that is initiated when the product of this gene binds a receptor on the surface of the monocyte cell. Thus, polynucleotides and polypeptides have uses which include, but are not limited to, activating monocyte cells. Furthermore, when tested against
20 U937 Myeloid cell lines, supernatants removed from cells containing this gene activated the GAS assay. Thus, it is likely that this gene activates myeloid cells through the Jak-STAT signal transduction pathway. The gamma activating sequence (GAS) is a promoter element found upstream of many genes which are involved in the Jak-STAT pathway. The Jak-STAT pathway is a large, signal transduction pathway involved in
25 the differentiation and proliferation of cells. Therefore, activation of the Jak-STAT pathway, reflected by the binding of the GAS element, can be used to indicate proteins involved in the proliferation and differentiation of cells.

This gene is expressed primarily in infant brain, and to a lesser extent in adult cerebellum and frontal cortex.

30 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, neurological disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential
35 identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the nervous system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell

types (e.g., cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:200 as residues: Glu-25 to Lys-33, Glu-115 to Lys-120.

The tissue distribution primarily in brain and homology to Slit, a gene involved in axon pathway development, indicates polynucleotides and polypeptides corresponding to this gene are useful for the treatment/detection of neurodegenerative disease states and behavioural disorders such as Alzheimers Disease, Parkinsons Disease, spinal cord injury, brain injuries, crushed (optic) nerve, amyotrophic lateral sclerosis, diabetes caused nerve damage, strokes, epilepsy, multiple sclerosis, paraplegia retinal degeneration, Huntingtons Disease, facial nerve damage, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder and panic disorder. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:86 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 2660 of SEQ ID NO:86, b is an integer of 15 to 2674, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:86, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 77

The translation product of this gene shares sequence homology with human endothelial cell multimerin, which is a secreted protein that binds to the extracellular matrix and is thought to be involved in hemostasis. Multimerin is a factor V/Va-binding protein and may function as a carrier protein for platelet factor V (J. Biol Chem 1995 Aug 4;270(31):18246-51). Contact of cells with supernatant expressing the product of this gene increases the permeability of THP-1 Monocyte cells to calcium. Thus, it is

likely that the product of this gene is involved in a signal transduction pathway that is initiated when the product of this gene binds a receptor on the surface of the monocyte cell. Thus, polynucleotides and polypeptides have uses which include, but are not limited to, activating monocyte cells.

5 This gene is expressed primarily in a variety of hematopoietic cells including T-cells, dendritic cells and B-cells as well as cells and tissues of epithelial and endothelial origin including healing wounds and keratinocytes, as well as placenta.

 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a
10 biological sample and for diagnosis of diseases and conditions which include, but are not limited to, acute internal injury, blood clotting disorders and other disorders of hemostasis. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s)
15 or cell type(s). For a number of disorders of the above tissues or cells, particularly of the hemostasis, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., endothelial, immune, cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in
20 healthy tissue or bodily fluid from an individual not having the disorder.

 Preferred epitopes include those comprising a sequence shown in SEQ ID NO:201 as residues: Ala-43 to Trp-57, Ser-81 to Gly-88, Tyr-125 to Asp-134, Pro-141 to Gly-154, Val-172 to Glu-178, Lys-296 to Gly-305, Leu-307 to Arg-314, Thr-335 to His-341.

25 The tissue distribution in endothelial tissues, and the homology to human endothelial cell multimerin, indicates polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and/or treatment of disorders involving the vasculature. Elevated expression of this gene product by endothelial cells indicates that it may play vital roles in the regulation of endothelial cell function; secretion;
30 proliferation; or angiogenesis. Furthermore, the tissue distribution indicates polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and/or treatment of disorders of the placenta. Specific expression within the placenta indicates that this gene product may play a role in the proper establishment and maintenance of placental function. Alternately, this gene product may be produced by
35 the placenta and then transported to the embryo, where it may play a crucial role in the development and/or survival of the developing embryo or fetus. Expression of this gene product in a vascular-rich tissue such as the placenta also indicates that this gene

product may be produced more generally in endothelial cells or within the circulation. In such instances, it may play more generalized roles in vascular function, such as in angiogenesis. It may also be produced in the vasculature and have effects on other cells within the circulation, such as hematopoietic cells, as supported by the biological activity data mentioned previously. It may serve to promote the proliferation, survival, activation, and/or differentiation of hematopoietic cells, as well as other cells throughout the body. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:87 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1622 of SEQ ID NO:87, b is an integer of 15 to 1636, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:87, and where b is greater than or equal to a + 14.

20

FEATURES OF PROTEIN ENCODED BY GENE NO: 78

In specific embodiments, polypeptides of the invention comprise the following amino acid sequence:

25 HYXSTPGRVPVRQFAAASTSGGPWVPGGXLEAPFQVAPSLSHSTPVFPGLI
(SEQ ID NO:616). Polynucleotides encoding these polypeptides are also encompassed by the invention.

This gene is expressed primarily in osteoblasts.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, degenerative conditions of the bone including arthritis and osteoporosis. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the skeletal system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., skeletal, cancerous and wounded

tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

- 5 Preferred epitopes include those comprising a sequence shown in SEQ ID NO:202 as residues: Thr-45 to Cys-50, Met-55 to Pro-60.

The tissue distribution in osteoblasts indicates polynucleotides and polypeptides corresponding to this gene are useful for treating degenerative conditions of the bone mediated by alterations in the activity ratio of osteoblasts and osteoclasts. Furthermore, elevated levels of expression of this gene product in osteoblastoma indicates that it may play a role in the survival, proliferation, and/or growth of osteoblasts. Therefore, it may be useful in influencing bone mass in such conditions as osteoporosis. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

- 15 Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:88 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome.
- 20 Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1625 of SEQ ID NO:88, b is an integer of 15 to 1639, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:88, and where b is greater than or equal to a + 14.

25

FEATURES OF PROTEIN ENCODED BY GENE NO: 79

- The gene encoding the disclosed cDNA is thought to reside on chromosome 17.
- 30 Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 17. In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: ARGKYESAQPGGTQPEPGLGAR (SEQ ID NO:617). Polynucleotides encoding these polypeptides are also encompassed by the invention.

- 35 This gene is expressed primarily in pituitary, cerebellum and kidney and to a lesser extent in a range of fetal tissues including lung, heart and spleen.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, metabolic, neurological, and renal disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the endocrine, renal and nervous systems expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., kidney, fetal, brain, cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:203 as residues: Pro-29 to Gly-34, Gln-79 to Arg-84, Arg-146 to Arg-152, Ser-183 to Ser-193, Gly-233 to His-241, Tyr-265 to Pro-278, Thr-304 to Arg-320, Leu-328 to Gly-333, Glu-385 to Arg-399.

The high expression of a secreted gene in the pituitary indicates a role for this gene or gene product in the treatment/detection of metabolic disorders associated with the endocrine system, such as growth and developmental defects. Expression in the cerebellum indicates a role in the treatment/detection of neurodegenerative disease states and behavioural disorders such as Alzheimers Disease, Parkinsons Disease, Huntingtons Disease, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder and panic disorder. Expression in the kidney indicates a role in the treatment/detection of renal disorders such as kidney failure, Wilms Tumor and kidney stones, as well as nephritis, renal tubular acidosis, proteinuria, pyuria, edema, pyelonephritis, hydronephritis, nephrotic syndrome, crush syndrome, glomerulonephritis, hematuria, renal colic and congenital kidney abnormalities such as horseshoe kidney, polycystic kidney, and Falconi's syndrome. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. .

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:89 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more

polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1846 of SEQ ID NO:89, b is an integer of 15 to 1860, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:89, and where b is greater than or equal to a + 14.

5

FEATURES OF PROTEIN ENCODED BY GENE NO: 80

The translation product of this gene shares sequence homology with ras-related proteins in rats which is thought to be involved in cellular signaling.

This gene is expressed primarily in T-cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune system disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., immune, cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:204 as residues: Met-40 to Thr-46, Ala-57 to Glu-64, Ser-85 to Leu-91.

The tissue distribution in immune system tissues indicates polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the gene or protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. In addition, this gene product may have commercial utility

in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Expression of this gene product in T cells also strongly indicates a role for this protein in immune function and immune surveillance. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:90 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 825 of SEQ ID NO:90, b is an integer of 15 to 839, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:90, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 81

20

When tested against U937 Myeloid cell lines, supernatants removed from cells containing this gene activated the GAS assay. Thus, it is likely that this gene activates myeloid cells through the Jak-STAT signal transduction pathway. The gamma activating sequence (GAS) is a promoter element found upstream of many genes which are involved in the Jak-STAT pathway. The Jak-STAT pathway is a large, signal transduction pathway involved in the differentiation and proliferation of cells. Therefore, activation of the Jak-STAT pathway, reflected by the binding of the GAS element, can be used to indicate proteins involved in the proliferation and differentiation of cells.

30

This gene is expressed primarily in fetal liver, osteoclastoma and neutrophils.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, diseases of the bone, haemopoietic system and cancer. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune,

haemopoietic and bone, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., immune, skeletal, cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having
5 such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in fetal liver, osteoclastoma and neutrophils indicates polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of diseases of the bone, haemopoietic and immune systems, as well as
10 cancer. Furthermore, elevated levels of expression of this gene product in osteoclastoma indicates that it may play a role in the survival, proliferation, and/or growth of osteoclasts. Therefore, it may be useful in influencing bone mass in such conditions as osteoporosis. More generally, as evidenced by expression in fetal liver/spleen, as well as the biological activity data, this gene may play a role in the
15 survival, proliferation, and/or differentiation of hematopoietic cells in general, and may be of use in the augmentation of the numbers of stem cells and committed progenitors. Expression of this gene product in neutrophils also indicates that it may play a role in mediating responses to infection and controlling immunological responses, such as those that occur during immune surveillance. Protein, as well as, antibodies directed
20 against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:91 and may have been publicly available prior to conception of the present
25 invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1131 of SEQ ID NO:91, b is an integer of 15
30 to 1145, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:91, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 82

35

This gene is expressed primarily in endometrial stromal cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, female infertility. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the reproductive system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., reproductive, cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:206 as residues: Gln-26 to Asn-51.

The tissue distribution in endometrial cells indicates polynucleotides and polypeptides corresponding to this gene are useful for treating female infertility. The protein product is likely involved in preparation of the endometrium of implantation and could be administered either topically or orally. Alternatively, this gene could be transfected in gene-replacement treatments into the cells of the endometrium and the protein products could be produced. Similarly, these treatments could be performed during artificial insemination for the purpose of increasing the likelihood of implantation and development of a healthy embryo. In both cases this gene or its gene product could be administered at later stages of pregnancy to promote healthy development of the endometrium. Protein, as well as, antibodies directed against the protein may show utility as a tissue-specific marker and/or immunotherapy target for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:92 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 2036 of SEQ ID NO:92, b is an integer of 15 to 2050, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:92, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 83

5 The gene encoding the disclosed cDNA is thought to reside on chromosome 19. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 19.

 This gene is expressed primarily in tissues of the central nervous system (predominantly the cerebellum) and immune system (predominantly the tonsils).

10 Nucleic acids of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, disorders of the immune system and CNS. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the
15 tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system and neurological system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., immune, neurological, cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell
20 sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

 Preferred epitopes include those comprising a sequence shown in SEQ ID NO:207 as residues: Pro-43 to Leu-49, Pro-61 to Gly-66, Ser-71 to Ser-83.

25 The tissue distribution in the immune system indicates polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Expression of this gene product in tonsils indicates a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This
30 gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the gene or protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed
35 tissues. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. In

addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above
5 listed tissues. Furthermore, the tissue distribution in the central nervous system indicates polynucleotides and polypeptides corresponding to this gene are useful for the detection/treatment of neurodegenerative disease states and behavioural disorders such as Alzheimers Disease, Parkinsons Disease, Huntingtons Disease, Tourette Syndrome, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic
10 disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, the gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the developing embryo, or sexually-linked disorders. Protein, as well as, antibodies directed against the protein may show utility
15 as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:93 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the
20 scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1159 of SEQ ID NO:93, b is an integer of 15 to 1173, where both a and b correspond to the positions of nucleotide residues shown
25 in SEQ ID NO:93, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 84

30 This gene is expressed primarily in testes.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, reproductive disorders and testes diseases. Similarly, polypeptides and
35 antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the testes, expression of this gene at

significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., testes, cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:208 as residues: Lys-28 to His-35, Asn-58 to Gly-64, Thr-80 to Asn-86, Pro-96 to Glu-111, Pro-124 to Phe-133.

The tissue distribution indicates polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of testes disorders. Furthermore, the tissue distribution in testes indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of conditions concerning proper testicular function (e.g. endocrine function, sperm maturation), as well as cancer. Therefore, this gene product is useful in the treatment of male infertility and/or impotence. This gene product is also useful in assays designed to identify binding agents, as such agents (antagonists) are useful as male contraceptive agents. Similarly, the protein is believed to be useful in the treatment and/or diagnosis of testicular cancer. The testes are also a site of active gene expression of transcripts that may be expressed, particularly at low levels, in other tissues of the body. Therefore, this gene product may be expressed in other specific tissues or organs where it may play related functional roles in other processes, such as hematopoiesis, inflammation, bone formation, and kidney function, to name a few possible target indications. Protein, as well as, antibodies directed against the protein may show utility as a tissue-specific marker and/or immunotherapy target for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:94 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 808 of SEQ ID NO:94, b is an integer of 15 to 822, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:94, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 85

5 This gene is expressed primarily in infant brain, bone marrow and activated T-cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, developmental, immune and hematopoietic disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune, hematopoietic and developmental systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., immune, developmental, cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

20 Preferred epitopes include those comprising a sequence shown in SEQ ID NO:209 as residues: Asn-23 to Val-37.

Expression in infant brain indicates polynucleotides and polypeptides corresponding to this gene are useful for the detection/treatment of mental retardation and other developmental disorders in addition to neurodegenerative disease states and behavioural disorders such as Alzheimers Disease, Parkinsons Disease, Huntingtons Disease, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder and panic disorder. Expression of the gene in bone marrow and in B-cells indicates a role in the treatment and/or detection of immune disorders such as arthritis, asthma, immunodeficiency diseases and leukemia. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:95 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more

polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1063 of SEQ ID NO:95, b is an integer of 15 to 1077, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:95, and where b is greater than or equal to a + 14.

5

FEATURES OF PROTEIN ENCODED BY GENE NO: 86

The gene encoding the disclosed cDNA is thought to reside on chromosome 1. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 1. In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: SCGSSRRSAKRSLTLKLIDFSHRI (SEQ ID NO:618). Polynucleotides encoding these polypeptides are also encompassed by the invention.

This gene is expressed primarily in infant brain, fetal liver and fetal spleen, and to a lesser extent in macrophages, T-cells, erythroid cells and myeloid progenitor cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, neurological, developmental and immune disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the nervous, immune and developmental systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., immune, neurological, developing, cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:210 as residues: Val-34 to Leu-48, Val-51 to Gly-67, Lys-74 to Asp-81, Thr-93 to Glu-98, Ser-138 to His-149, Ala-186 to Gln-201, Pro-257 to Arg-271.

Expression in infant brain indicates polynucleotides and polypeptides corresponding to this gene are useful for the detection/treatment of mental retardation and other developmental disorders in addition to neurodegenerative disease states and behavioural disorders such as Alzheimers Disease, Parkinsons Disease, Huntingtons

Disease, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder and panic disorder. Its distribution in fetal liver and fetal spleen indicates that this gene may play a role in the development of the hematopoietic and immune systems and that it may play a role in the treatment/detection of immune system disorders such as leukemia, arthritis and asthma. Protein, as well as, antibodies directed against the protein may show utility as a tissue-specific marker and/or immunotherapy target for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:96 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 2078 of SEQ ID NO:96, b is an integer of 15 to 2092, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:96, and where b is greater than or equal to a + 14.

20 FEATURES OF PROTEIN ENCODED BY GENE NO: 87

This gene is expressed primarily in hippocampus, and to a lesser extent in fetal heart.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, any of a variety of brain disorders including epilepsy, stroke, palsy, and mood disorders including unipolar and bipolar depression. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the central nervous system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., brain, heart, cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in hippocampus indicates polynucleotides and polypeptides corresponding to this gene are useful for the detection/treatment of neurodegenerative disease states and behavioural disorders such as Alzheimers Disease, Parkinsons Disease, Huntingtons Disease, Tourette Syndrome, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic disorder, learning disabilities, epilepsy, stroke, palsy, and mood disorders including unipolar and bipolar depression, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, the gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the developing embryo, or sexually-linked disorders. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:97 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1338 of SEQ ID NO:97, b is an integer of 15 to 1352, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:97, and where b is greater than or equal to a + 14.

25 FEATURES OF PROTEIN ENCODED BY GENE NO: 88

The translation product of this gene shares sequence homology with a C. elegans protein (coded for by C. elegans cDNA yk112f3.5). The gene encoding the disclosed cDNA is thought to reside on chromosome 3. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 3. In specific embodiments, polypeptides of the invention comprise the following amino acid sequence:

HYFLRTVSGLSVVPVSLRCCMCPPTGPAPATAHSPFDPPALPIQFEYQQA
(SEQ ID NO:619), QLEAEIENLSWKVERADSYDRGDLENQMHIAEQRRRT
35 LLKDFHDT (SEQ ID NO:620), VPSLRCCMCPPTGPAPATAHS (SEQ ID
NO:621), and/or SWKVERADSYDRGDLENQMHIAEQR (SEQ ID NO:622).
Polynucleotides encoding these polypeptides are also encompassed by the invention.

This gene is expressed primarily in fetal liver and spleen.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, congenital disorders of the liver and spleen. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the hepatic system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., immune, hepatic, cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in fetal liver and spleen indicates polynucleotides and polypeptides corresponding to this gene are useful for the detection and treatment of liver disorders and cancers (e.g. hepatoblastoma, jaundice, hepatitis, liver metabolic diseases and conditions that are attributable to the differentiation of hepatocyte progenitor cells). More generally, as evidenced by expression in fetal liver/spleen, this gene may play a role in the survival, proliferation, and/or differentiation of hematopoietic cells in general, and may be of use in augmentation of the numbers of stem cells and committed progenitors. Expression of this gene product in primary dendritic cells also indicates that it may play a role in mediating responses to infection and controlling immunological responses, such as those that occur during immune surveillance. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:98 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 899 of SEQ ID NO:98, b is an integer of 15 to 913, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:98, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 89

5 This gene is expressed primarily in neutrophils.

 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, acute immunological disorders such as inflammation. Similarly,
10 polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., immune, cancerous and wounded tissues) or bodily
15 fluids (e.g., serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

 The tissue distribution in neutrophils indicates polynucleotides and polypeptides
20 corresponding to this gene are useful for treating an acute inflammatory response. Furthermore, this gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the gene or protein, as well as, antibodies
25 directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. In addition, this gene product may have commercial utility in the expansion
30 of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Expression of this gene product in neutrophils also strongly indicates a role for this protein in immune function and immune surveillance. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed
35 tissues.

 Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ

ID NO:99 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome.

Accordingly, preferably excluded from the present invention are one or more

- 5 polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 707 of SEQ ID NO:99, b is an integer of 15 to 721, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:99, and where b is greater than or equal to a + 14.

10

FEATURES OF PROTEIN ENCODED BY GENE NO: 90

- When tested against U937 Myeloid cell lines, supernatants removed from cells containing this gene activated the GAS assay. Thus, it is likely that this gene activates
15 myeloid cells through the Jak-STAT signal transduction pathway. The gamma activating sequence (GAS) is a promoter element found upstream of many genes which are involved in the Jak-STAT pathway. The Jak-STAT pathway is a large, signal transduction pathway involved in the differentiation and proliferation of cells. Therefore, activation of the Jak-STAT pathway, reflected by the binding of the GAS
20 element, can be used to indicate proteins involved in the proliferation and differentiation of cells.

This gene is expressed primarily in fetal liver and fetal spleen.

- Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a
25 biological sample and for diagnosis of diseases and conditions which include, but are not limited to, developmental and/or immune disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and developmental systems,
30 expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., liver, spleen, cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily
35 fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:214 as residues: His-23 to Leu-31, His-33 to Pro-41.

The distribution of this gene in fetal liver and fetal spleen and the biological activity data indicates it may play a role in the development of the immune and hematopoietic systems. It may, therefore, play a role in the treatment and/or detection of immune and/or hematopoietic disorders including leukemia, arthritis and asthma.

5 Furthermore, this gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the gene or protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Therefore it may be also used as an agent for
10 immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the
15 differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ
20 ID NO:100 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of
25 a-b, where a is any integer between 1 to 631 of SEQ ID NO:100, b is an integer of 15 to 645, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:100, and where b is greater than or equal to a + 14.

30 FEATURES OF PROTEIN ENCODED BY GENE NO: 91

This gene is expressed primarily in brain and osteoclastoma to a lesser extent in placenta.

Therefore, polynucleotides and polypeptides of the invention are useful as
35 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, neurological, bone and reproductive disorders. Similarly, polypeptides

and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the nervous, bone and reproductive systems, expression of this gene at significantly higher or lower levels
5 may be routinely detected in certain tissues or cell types (e.g., skeletal, brain, cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

10 Preferred epitopes include those comprising a sequence shown in SEQ ID NO:215 as residues: Phe-47 to Cys-54.

Expression of this gene in brain indicates polynucleotides and polypeptides corresponding to this gene are useful for the detection/treatment of neurodegenerative disease states and behavioural disorders such as Alzheimers Disease, Parkinsons
15 Disease, Huntingtons Disease, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder and panic disorder. Furthermore, expression in osteoclastoma indicates a role in the treatment and/or detection of bone damage such as fractures and dislocations. Elevated levels of expression of this gene product in osteoclastoma indicates that it may play a role in the survival, proliferation, and/or growth of
20 osteoclasts. Therefore, it may be useful in influencing bone mass in such conditions as osteoporosis. Expression in the placenta indicates a role in the treatment and/or detection of pregnancy disorders such as miscarriage, birth defects, premature birth, in addition to disorders such as placenta previa and placentitis. Specific expression within the placenta indicates that this gene product may play a role in the proper establishment
25 and maintenance of placental function. Alternately, this gene product may be produced by the placenta and then transported to the embryo, where it may play a crucial role in the development and/or survival of the developing embryo or fetus. Expression of this gene product in a vascular-rich tissue such as the placenta also indicates that this gene product may be produced more generally in endothelial cells or within the circulation. In
30 such instances, it may play more generalized roles in vascular function, such as in angiogenesis. It may also be produced in the vasculature and have effects on other cells within the circulation, such as hematopoietic cells. It may serve to promote the proliferation, survival, activation, and/or differentiation of hematopoietic cells, as well as other cells throughout the body. Protein, as well as, antibodies directed against the
35 protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:101 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 549 of SEQ ID NO:101, b is an integer of 15 to 563, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:101, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 92

- 15 This gene is expressed primarily in T-cells, bone marrow, fetal liver/spleen and to a lesser extent in adipocytes, kidney, melanocytes and stimulated fibroblasts.
- Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, hematopoietic disease characterized by alterations in T cells. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., immune, cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.
- 30 The tissue distribution in immune tissues indicates polynucleotides and polypeptides corresponding to this gene are useful for treating autoimmune diseases or proliferative disorders of the developing immune system. This gene product is primarily expressed in hematopoietic cells and tissues, suggesting that it plays a role in the survival, proliferation, and/or differentiation of hematopoietic lineages. This is particularly supported by the expression of this gene product in fetal liver and bone marrow, the two primary sites of definitive hematopoiesis. Expression of this gene product in T cells also strongly indicates a role for this protein in immune function and
- 35

immune surveillance. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:102 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1310 of SEQ ID NO:102, b is an integer of 15 to 1324, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:102, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 93

The gene encoding the disclosed cDNA is thought to reside on chromosome 4. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 4.

This gene is expressed primarily in fetal tissues including fetal liver/spleen and to a lesser extent in lung, bone marrow, adrenal gland tumor and in the Ntera2 cell line.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, cancers. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the adrenal gland or lungs, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., developing tissues, cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in fetal and developing tissues indicates polynucleotides and polypeptides corresponding to this gene are useful for treating tumors formed by

poorly differentiated cells, as well as tumors of other tissues where expression has been observed. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available
5 and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:103 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more
10 polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1717 of SEQ ID NO:103, b is an integer of 15 to 1731, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:103, and where b is greater than or equal to a + 14.

15

FEATURES OF PROTEIN ENCODED BY GENE NO: 94

This gene is expressed primarily in the prostate derived cell line PC3 and fetal liver/spleen.

20 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, prostatic hypertrophy or prostate cancer. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes
25 for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the glandular tissues, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., prostate, immune, cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell
30 sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:218 as residues: Leu-26 to Ser-33.

35 The tissue distribution indicates polynucleotides and polypeptides corresponding to this gene are useful for treating diseases of the prostate including prostatic tumors or benign prostatic hypertrophy. Protein, as well as, antibodies

directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:104 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1452 of SEQ ID NO:104, b is an integer of 15 to 1466, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:104, and where b is greater than or equal to a + 14.

15 FEATURES OF PROTEIN ENCODED BY GENE NO: 95

This gene is expressed primarily in small intestine, and to a lesser extent in breast tissue.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, diseases of the small intestine. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the small intestine, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., small intestine, cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:219 as residues: Glu-37 to Gly-45.

The tissue distribution indicates polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of diseases involving the small intestine, such as cancer of the small intestine or other tissues where expression has been indicated. Protein, as well as, antibodies directed against the

protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:105 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1289 of SEQ ID NO:105, b is an integer of 15 to 1303, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:105, and where b is greater than or equal to a + 14.

15 FEATURES OF PROTEIN ENCODED BY GENE NO: 96

This gene is expressed primarily in fast-growing tissues such as tumor and fetal tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, growth disorders such as tumorigenesis and growth retardation. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the fast-growing tissues, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., rapidly proliferating cells, cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:220 as residues: Phe-32 to Cys-37.

The tissue distribution in rapidly-proliferating tissues and cells indicates polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis, treatment, and/or prevention of growth disorders. Furthermore, the tissue distribution indicates polynucleotides and polypeptides corresponding to this gene are

useful for the diagnosis and treatment of cancer and other proliferative disorders. Expression within embryonic tissue and other cellular sources marked by proliferating cells indicates that this protein may play a role in the regulation of cellular division. Similarly, embryonic development also involves decisions involving cell differentiation and/or apoptosis in pattern formation. Thus, this protein may also be involved in apoptosis or tissue differentiation and could again be useful in cancer therapy. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:106 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1502 of SEQ ID NO:106, b is an integer of 15 to 1516, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:106, and where b is greater than or equal to a + 14.

20

FEATURES OF PROTEIN ENCODED BY GENE NO: 97

When tested against Jurkat T-cells and U937 Myeloid cell lines, supernatants removed from cells containing this gene activated the GAS assay. Thus, it is likely that this gene activates both T-cells and myeloid cells through the Jak-STAT signal transduction pathway. The gamma activating sequence (GAS) is a promoter element found upstream of many genes which are involved in the Jak-STAT pathway. The Jak-STAT pathway is a large, signal transduction pathway involved in the differentiation and proliferation of cells. Therefore, activation of the Jak-STAT pathway, reflected by the binding of the GAS element, can be used to indicate proteins involved in the proliferation and differentiation of cells.

This gene is expressed primarily in placenta, and to a lesser extent in the endometrium.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, pregnancy disorders. Similarly, polypeptides and antibodies directed to

these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the female reproductive system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., placental, reproductive, cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Expression of this gene in the placenta and endometrium indicates a role in the treatment and/or detection of pregnancy disorders such as miscarriage, birth defects, premature birth, in addition to disorders such as endometriosis, placenta previa and placentitis. Furthermore, the tissue distribution indicates polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and/or treatment of disorders of the placenta. Specific expression within the placenta indicates that this gene product may play a role in the proper establishment and maintenance of placental function. Alternately, this gene product may be produced by the placenta and then transported to the embryo, where it may play a crucial role in the development and/or survival of the developing embryo or fetus. Expression of this gene product in a vascular-rich tissue such as the placenta also indicates that this gene product may be produced more generally in endothelial cells or within the circulation. In such instances, it may play more generalized roles in vascular function, such as in angiogenesis. It may also be produced in the vasculature and have effects on other cells within the circulation, such as hematopoietic cells. It may serve to promote the proliferation, survival, activation, and/or differentiation of hematopoietic cells, as well as other cells throughout the body. Alternatively, the tissue distribution indicates polynucleotides and polypeptides corresponding to this gene are useful for treating female infertility. The protein product is likely involved in preparation of the endometrium of implantation and could be administered either topically or orally. Alternatively, this gene could be transfected in gene-replacement treatments into the cells of the endometrium and the protein products could be produced. Similarly, these treatments could be performed during artificial insemination for the purpose of increasing the likelihood of implantation and development of a healthy embryo. In both cases this gene or its gene product could be administered at later stages of pregnancy to promote healthy development of the endometrium. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:107 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1675 of SEQ ID NO:107, b is an integer of 15 to 1689, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:107, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 98

The gene encoding the disclosed cDNA is thought to reside on chromosome 5. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 5. Recently another group gened and sequenced this gene, calling it MDC-3.13 isoform 1 (Genbank Accession Number: g3860095), which is believed to be a cellular factor involved in the differentiation of dendritic cells. In specific embodiments, polypeptides of the invention comprise the following amino acid sequence:

HEAWLRSAGTREPPREQRTRRRQTAQLALQVPAPSRTPPMATDVFNSKNLAVX
 AQKKILGKMVSKSIATTLIDDSSEVLDELRYRTREYTONKKEAEKIIKNLIKTVI
 KLAILYRNNQFNQDELALMEKFKKKVHQLAMTVVSFHQVDYTFDRNVLSRLL
 NECREMLHQIQRHLTAKSHGRVNNVFDHFSDCEFLAALYNPFGNFKPHLQKL
 CDGINKMLDEENI (SEQ ID NO:623), HEAWLRSAGTREPPREQRTRRRQTAQLA
 LQVPAPSRTPPMATDVFNSKNLAV (SEQ ID NO:624), XAQQKILGKMVSKSIAT
 TLIDDSSEVLDELRYRTREYTONKKEAEKII (SEQ ID NO:625), KNLIKTVIKLA
 ILYRNNQFNQDELALMEKFKKKVHQLAMTVVSFHQVDYTF (SEQ ID NO:626),
 DRNVLSRLLNECREMLHQIQRHLTAKSHGRVNNVFDHFSDCEFLAALYNPF
 (SEQ ID NO:627), and/or GNFKPHLQKLCDGINKMLDEENI (SEQ ID NO:628).
 Polynucleotides encoding these polypeptides are also encompassed by the invention.

This gene is expressed primarily in placenta, spleen from CLL patients and various T cell libraries, and to a lesser extent in lung, bone marrow, neutrophil, osteoclastoma, and lymphoma tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a

biological sample and for diagnosis of diseases and conditions which include, but are not limited to, diseases of the blood particularly diseases afflicting T cells and tumors of blood cells. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the hematopoietic system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., placental, immune, vascular, cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in immune tissues indicates polynucleotides and polypeptides corresponding to this gene are useful for treating diseases of the blood including leukemias, lymphomas and diseases that alter T-cell function or proliferation. Furthermore, the tissue distribution in placenta indicates polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and/or treatment of disorders of the placenta. Specific expression within the placenta indicates that this gene product may play a role in the proper establishment and maintenance of placental function. Alternately, this gene product may be produced by the placenta and then transported to the embryo, where it may play a crucial role in the development and/or survival of the developing embryo or fetus. Expression of this gene product in a vascular-rich tissue such as the placenta also indicates that this gene product may be produced more generally in endothelial cells or within the circulation. In such instances, it may play more generalized roles in vascular function, such as in angiogenesis. It may also be produced in the vasculature and have effects on other cells within the circulation, such as hematopoietic cells. It may serve to promote the proliferation, survival, activation, and/or differentiation of hematopoietic cells, as well as other cells throughout the body. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:108 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more

polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1929 of SEQ ID NO:108, b is an integer of 15 to 1943, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:108, and where b is greater than or equal to a + 14.

5

FEATURES OF PROTEIN ENCODED BY GENE NO: 99

This gene is expressed primarily in rejected kidney, placenta, and melanocytes.

10 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, acute or chronic renal failure. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for
15 differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the renal system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., renal, placental, cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample
20 taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:223 as residues: Thr-41 to Pro-47.

25 The tissue distribution in kidney indicates polynucleotides and polypeptides corresponding to this gene are useful for treating diseases of the kidney, including renal failure of either an acute or chronic nature, as well as nephritis, renal tubular acidosis, proteinuria, pyuria, edema, pyelonephritis, hydronephritis, nephrotic syndrome, crush syndrome, glomerulonephritis, hematuria, renal colic and kidney stones, in addition to
30 Wilms Tumor Disease, and congenital kidney abnormalities such as horseshoe kidney, polycystic kidney, and Falconi's syndrome. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

35 Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:109 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the

scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1580 of SEQ ID NO:109, b is an integer of 15
5 to 1594, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:109, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 100

10

When tested against U937 Myeloid cell lines, supernatants removed from cells containing this gene activated the GAS assay. Thus, it is likely that this gene activates myeloid cells through the Jak-STAT signal transduction pathway. The gamma activating sequence (GAS) is a promoter element found upstream of many genes which
15 are involved in the Jak-STAT pathway. The Jak-STAT pathway is a large, signal transduction pathway involved in the differentiation and proliferation of cells. Therefore, activation of the Jak-STAT pathway, reflected by the binding of the GAS element, can be used to indicate proteins involved in the proliferation and differentiation of cells.

20

This gene is expressed primarily in spinal cord, and to a lesser extent in melanocytes and fetal spleen/liver.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are
25 not limited to, central nervous system diseases. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the central nervous system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues
30 or cell types (e.g., central nervous system, cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

35

The tissue distribution in spinal cord tissue indicates polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of central nervous system disorders, such as Alzheimers Disease, Parkinsons Disease,

Huntingtons Disease, Tourette Syndrome, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, the gene or gene product may also play a role in the
5 treatment and/or detection of developmental disorders associated with the developing embryo, or sexually-linked disorders. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available
10 and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:110 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more
15 polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1728 of SEQ ID NO:110, b is an integer of 15 to 1742, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:110, and where b is greater than or equal to a + 14.

20

FEATURES OF PROTEIN ENCODED BY GENE NO: 101

This gene is expressed primarily in breast and dendritic cells.

Therefore, polynucleotides and polypeptides of the invention are useful as
25 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, breast related disorders and inflammatory diseases. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a
30 number of disorders of the above tissues or cells, particularly of the breast tissue and dendritic cells, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., breast, immune, cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder,
35 relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in breast indicates polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of breast related diseases and inflammatory disorders. Furthermore, the tissue distribution in breast indicates polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and intervention of breast tumors, in addition to other tumors where expression has been indicated. Protein, as well as, antibodies directed against the protein may show utility as a tissue-specific marker and/or immunotherapy target for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:111 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1487 of SEQ ID NO:111, b is an integer of 15 to 1501, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:111, and where b is greater than or equal to a + 14.

20

FEATURES OF PROTEIN ENCODED BY GENE NO: 102

When tested against sensory neuron cell lines, supernatants removed from cells containing this gene activated the EGR1 assay. Thus, it is likely that this gene activates neuronal cells through a signal transduction pathway. Early growth response 1 (EGR1) is a promoter associated with certain genes that induces various tissues and cell types upon activation, leading the cells to undergo differentiation and proliferation. Furthermore, when tested against both Jurkat T-cells and U937 Myeloid cell lines, supernatants removed from cells containing this gene activated the GAS assay. Thus, it is likely that this gene activates both T-cells and myeloid cells through the Jak-STAT signal transduction pathway. The gamma activating sequence (GAS) is a promoter element found upstream of many genes which are involved in the Jak-STAT pathway. The Jak-STAT pathway is a large, signal transduction pathway involved in the differentiation and proliferation of cells. Therefore, activation of the Jak-STAT pathway, reflected by the binding of the GAS element, can be used to indicate proteins involved in the proliferation and differentiation of cells.

This gene is expressed primarily in synovial sarcoma.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, synovial sarcoma. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the synovial sarcoma, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., synovium, cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in synovial sarcoma, and the biological activity data, suggest that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of synovial sarcoma. In general, the expression of this gene product in synovium indicates a role in the detection and treatment of disorders and conditions afflicting the skeletal system, in particular osteoporosis as well as disorders afflicting connective tissues (e.g. arthritis, trauma, tendonitis, chondromalacia and inflammation), such as in the diagnosis or treatment of various autoimmune disorders such as rheumatoid arthritis, lupus, scleroderma, and dermatomyositis as well as dwarfism, spinal deformation, and specific joint abnormalities as well as chondrodysplasias (ie. spondyloepiphyseal dysplasia congenita, familial osteoarthritis, Atelosteogenesis type II, metaphyseal chondrodysplasia type Schmid). Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:112 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 777 of SEQ ID NO:112, b is an integer of 15 to 791, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:112, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 103

This gene is expressed primarily in human tonsils.

Therefore, polynucleotides and polypeptides of the invention are useful as
5 reagents for differential identification of the tissue(s) or cell type(s) present in a
biological sample and for diagnosis of diseases and conditions which include, but are
not limited to, relating to inflammatory diseases such as tonsilitis, and immune system
disorders. Similarly, polypeptides and antibodies directed to these polypeptides are
useful in providing immunological probes for differential identification of the tissue(s)
10 or cell type(s). For a number of disorders of the above tissues or cells, particularly of
the immune system, expression of this gene at significantly higher or lower levels may
be routinely detected in certain tissues or cell types (e.g., immune, cancerous and
wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid and spinal
fluid) or another tissue or cell sample taken from an individual having such a disorder,
15 relative to the standard gene expression level, i.e., the expression level in healthy tissue
or bodily fluid from an individual not having the disorder.

The tissue distribution in tonsils indicates polynucleotides and polypeptides
corresponding to this gene are useful for the diagnosis, treatment, and/or prevention of
lymphoid tissue disorders such as tonsilitis. Furthermore, the tissue distribution
20 indicates polynucleotides and polypeptides corresponding to this gene are useful for the
diagnosis and treatment of a variety of immune system disorders. Expression of this
gene product in tonsils indicates a role in the regulation of the proliferation; survival;
differentiation; and/or activation of potentially all hematopoietic cell lineages, including
blood stem cells. This gene product may be involved in the regulation of cytokine
25 production, antigen presentation, or other processes that may also suggest a usefulness
in the treatment of cancer (e.g. by boosting immune responses). Since the gene is
expressed in cells of lymphoid origin, the gene or protein, as well as, antibodies
directed against the protein may show utility as a tumor marker and/or immunotherapy
targets for the above listed tissues. Therefore it may be also used as an agent for
30 immunological disorders including arthritis, asthma, immune deficiency diseases such
as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and
psoriasis. In addition, this gene product may have commercial utility in the expansion
of stem cells and committed progenitors of various blood lineages, and in the
differentiation and/or proliferation of various cell types. Protein, as well as, antibodies
35 directed against the protein may show utility as a tumor marker and/or immunotherapy
targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:113 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1623 of SEQ ID NO:113, b is an integer of 15 to 1637, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:113, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 104

This gene is expressed primarily in activated T-cells and prostate tissue. Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, T lymphocytes related diseases and inflammation of the prostate. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and reproductive systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., immune, reproductive, cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:228 as residues: Arg-24 to Trp-36.

The tissue distribution in immune system tissues and prostate tissue indicates polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of immune and reproductive disorders. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin,

the gene or protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Expression of this gene product in T cells also strongly indicates a role for this protein in immune function and immune surveillance. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:114 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1574 of SEQ ID NO:114, b is an integer of 15 to 1588, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:114, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 105

This gene is expressed primarily in human adult pulmonary tissue and infant brain.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, relating to the lung, neurological and immunological disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the respiratory, nervous, and immune systems expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., pulmonary, immune, nervous, cancerous and wounded tissues) or bodily fluids (e.g.,

serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

- 5 The tissue distribution in pulmonary tissue and infant brain tissue indicates polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment for disorders relating to the pulmonary system, the central nervous system, and the immune system. Furthermore, the tissue distribution in pulmonary tissue and fetal tissue indicates polynucleotides and polypeptides
- 10 corresponding to this gene are useful for the detection and treatment of disorders associated with developing lungs, particularly in premature infants where the lungs are the last tissues to develop. The tissue distribution indicates polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and intervention of lung tumors, since the gene may be involved in the regulation of cell division.
- 15 Additionally, the tissue distribution indicates polynucleotides and polypeptides corresponding to this gene are useful for the detection/treatment of neurodegenerative disease states and behavioural disorders such as Alzheimers Disease, Parkinsons Disease, Huntingtons Disease, Tourette Syndrome, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic disorder, learning disabilities, ALS,
- 20 psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, the gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the developing embryo, or sexually-linked disorders. Also, this gene product may be involved in the regulation of cytokine production, antigen presentation, or other
- 25 processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the gene or protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma,
- 30 immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a
- 35 tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ

ID NO:115 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome.

Accordingly, preferably excluded from the present invention are one or more
5 polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1912 of SEQ ID NO:115, b is an integer of 15 to 1926, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:115, and where b is greater than or equal to a + 14.

10

FEATURES OF PROTEIN ENCODED BY GENE NO: 106

The translation product of this gene shares sequence homology with the KIAA0132 gene product, and also shares homology to Drosophila melanogaster ring
15 canel protein. The gene encoding the disclosed cDNA is thought to reside on chromosome 1. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 1.

This gene is expressed primarily in infant brain and B-cell lymphoma.

Therefore, polynucleotides and polypeptides of the invention are useful as
20 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, relating to the central nervous system and B cell disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a
25 number of disorders of the above tissues or cells, particularly of the central nervous and immune systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., central nervous system, immune, cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having
30 such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:230 as residues: Thr-31 to Trp-42, Gly-49 to His-54, Gly-68 to Glu-75, Ser-77 to Trp-89, Met-142 to Gly-148.

35 The tissue distribution in infant brain and B-cell lymphomas indicates polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis, treatment, and/or prevention for central nervous system and immune

disorders. Furthermore, the tissue distribution indicates polynucleotides and polypeptides corresponding to this gene are useful for the detection/treatment of neurodegenerative disease states and behavioural disorders such as Alzheimers Disease, Parkinsons Disease, Huntingtons Disease, Tourette Syndrome, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, the gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the developing embryo, or sexually-linked disorders. Additionally, this gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the gene or protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:116 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1049 of SEQ ID NO:116, b is an integer of 15 to 1063, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:116, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 107

When tested against sensory neuron cell lines, supernatants removed from cells containing this gene activated the EGR1 assay. Thus, it is likely that this gene activates neuronal cells through a signal transduction pathway. Early growth response 1 (EGR1) is a promoter associated with certain genes that induces various tissues and cell types upon activation, leading the cells to undergo differentiation and proliferation. The gene encoding the disclosed cDNA is thought to reside on chromosome 1. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 1.

This gene is expressed primarily in human gall bladder.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, relating to gastrointestinal disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the gastrointestinal system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., gall bladder, cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:231 as residues: Pro-45 to Pro-51.

The tissue distribution in gall bladder indicates polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of gastrointestinal disorders. Protein, as well as, antibodies directed against the protein may show utility as a tissue-specific marker and/or immunotherapy target for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:117 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of

a-b, where a is any integer between 1 to 1601 of SEQ ID NO:117, b is an integer of 15 to 1615, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:117, and where b is greater than or equal to a + 14.

5

FEATURES OF PROTEIN ENCODED BY GENE NO: 108

This gene is expressed primarily in human whole brain.

Therefore, polynucleotides and polypeptides of the invention are useful as
10 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, neurodegenerative diseases. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of
15 the above tissues or cells, particularly of the central nervous and endocrine systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., brain, cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard
20 gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:232 as residues: Gln-58 to Asp-64, His-69 to Pro-76, Leu-101 to Glu-108.

The tissue distribution in brain tissue indicates polynucleotides and polypeptides
25 corresponding to this gene are useful for the diagnosis and treatment of the central nervous system and endocrine system disorders. Furthermore, the tissue distribution indicates polynucleotides and polypeptides corresponding to this gene are useful for the detection/treatment of neurodegenerative disease states and behavioural disorders such as Alzheimers Disease, Parkinsons Disease, Huntingtons Disease, Tourette Syndrome,
30 schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, the gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the developing embryo, or sexually-linked
35 disorders. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:118 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1207 of SEQ ID NO:118, b is an integer of 15 to 1221, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:118, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 109

The translation product of this gene shares sequence homology with human translation initiation factor eIF3 p40 subunit.

This gene is expressed primarily in human adipose, human fetal spleen, and dendritic cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, adipose, immune and nerve cell disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and nervous systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., adipose, immune, nervous, cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:233 as residues: Asn-27 to Ser-33, Gln-44 to Lys-50.

The tissue distribution fetal liver/spleen and dendritic cells indicates polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of immune and nerve cell disorders. Furthermore, the tissue distribution in adipose tissue indicates polynucleotides and polypeptides corresponding

to this gene are useful for the treatment of obesity and other metabolic and endocrine conditions or disorders. Additionally, the protein product of this gene may show utility in ameliorating conditions which occur secondary to aberrant fatty-acid metabolism (e.g. aberrant myelin sheath development), either directly or indirectly. Also, the tissue distribution indicates polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and/or treatment of hematopoietic disorders. This gene product is primarily expressed in hematopoietic cells and tissues, suggesting that it plays a role in the survival, proliferation, and/or differentiation of hematopoietic lineages. This is particularly supported by the expression of this gene product in fetal liver, which is a primary site of definitive hematopoiesis. Expression of this gene product in primary dendritic cells also strongly indicates a role for this protein in immune function and immune surveillance. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:119 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1135 of SEQ ID NO:119, b is an integer of 15 to 1149, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:119, and where b is greater than or equal to a + 14.

25

FEATURES OF PROTEIN ENCODED BY GENE NO: 110

The translation product of this gene shares sequence homology with Ig V-chain, which is thought to be important in immune function. When tested against Jurkat cell lines, supernatants removed from cells containing this gene activated the NF-kB transcription factor. Thus, it is likely that this gene activates Jurkat cells by activating a transcriptional factor found within these cells. Nuclear factor kB is a transcription factor activated by a wide variety of agents, leading to cell activation, differentiation, or apoptosis. Reporter constructs utilizing the NF-kB promoter element are used to screen supernatants for such activity.

35

This gene is expressed in human synovial sarcoma, infant brain INIB cells, macrophages (GM-CSF treated), human endometrial stromal cells-treated with estradiol, human pancreas tumor, hemangiopericytoma, human endometrial tumor, chronic lymphocytic leukemia and a human colon carcinoma (HCC) cell line.

5 Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, cancers such as human synovial sarcoma, human pancreas tumor, hemangiopericytoma, human endometrial tumor, chronic lymphocytic leukemia and
10 human colon carcinoma. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., immune,
15 cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Preferred epitopes include those comprising a sequence shown in SEQ ID
20 NO:234 as residues: Leu-21 to Ala-30, Ser-38 to Asp-47, Pro-87 to Asp-94, Leu-197 to Thr-204, Pro-256 to Ser-262, Thr-277 to Arg-282, Thr-293 to Trp-303.

The tissue distribution in numerous cancerous tissues, and the homology to Ig V-chain, as well as the biological activity data, indicates polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of
25 cancers, including human synovial sarcoma, human pancreas tumor, hemangiopericytoma, human endometrial tumor, chronic lymphocytic leukemia and human colon carcinoma, as well as other tissues where expression has been demonstrated. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

30 Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:120 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome.

35 Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1501 of SEQ ID NO:120, b is an integer of 15

to 1515, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:120, and where b is greater than or equal to a + 14.

Gene No.	cDNA Clone ID	ATCC Deposit Nr and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	First AA of Secreted Portion	Last AA of ORF
1	HFCCQ50	209463 11/14/97	Uni-ZAP XR	11	1271	1	1271	47	47	125	1	20	21	352
2	HTLAI54	209463 11/14/97	Uni-ZAP XR	12	1451	1	1451	125	125	126	1	22	23	157
3	HKABT24	209463 11/14/97	pCMVSPORT 2.0	13	2317	1	1809	66	66	127	1	22	23	553
4	HLWBF94	209463 11/14/97	pCMVSPORT 3.0	14	1472	1	1472	192	192	128	1	18	19	307
5	HFKEFF78	209463 11/14/97	Uni-ZAP XR	15	1016	1	961	92	92	129	1	18	19	166
6	HSYBG37	209463 11/14/97	pCMVSPORT 3.0	16	1239	1	1239	48	48	130	1	24	25	305
7	HTHCA77	209463 11/14/97	Uni-ZAP XR	17	1405	1	1405	160	160	131	1	24	25	219

Gene No.	cDNA Clone ID	ATCC Deposit Nr and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	First AA of Secreted Portion	Last AA of ORF
8	HNHZE51	209463 11/14/97	Uni-ZAP XR	18	1534	1	1534	106	106	132	1	32	33	98
9	HFIAX46	209463 11/14/97	pSport1	19	1233	1	1233	195	195	133	1	20	21	60
10	HFOXO72	209463 11/14/97	pSport1	20	1090	1	862	240	240	134	1	32	33	247
11	HODDW40	209463 11/14/97	Uni-ZAP XR	21	682	1	682	139	139	135	1	19	20	40
12	HSAWG42	209463 11/14/97	Uni-ZAP XR	22	770	1	770	31	31	136	1	27	28	74
13	HBMSK09	209463 11/14/97	Uni-ZAP XR	23	565	1	565	67	67	137	1	28	29	74
14	HDP AU16	209463 11/14/97	pCMV Sport 3.0	24	1356	1	1356	486	486	138	1	24	25	57
15	HFE BE12	209463 11/14/97	Uni-ZAP XR	25	617	3	617	99	99	139	1	22	23	173

Gene No.	cDNA Clone ID	ATCC Deposit Nr and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	First AA of Secreted Portion	Last AA of ORF
16	HFLNB64	209463 11/14/97	Uni-ZAP XR	26	648	1	648	62	62	140	1	39	40	45
16	HCESD11	209877 05/18/98	pBluescript	121	1025	1	1025	188	188	235	1	18	19	28
17	HS AWZ41	209463 11/14/97	Uni-ZAP XR	27	1388	1	1388	98	98	141	1	24	25	57
18	HNFJF07	209463 11/14/97	Uni-ZAP XR	28	616	1	616	86	86	142	1	21	22	66
19	HNGJO57	209463 11/14/97	Uni-ZAP XR	29	828	1	828	87	87	143	1	18	19	52
20	HE7TM22	209463 11/14/97	Uni-ZAP XR	30	581	1	581	70	70	144	1	22	23	65
21	HFRBR70	209463 11/14/97	Uni-ZAP XR	31	789	1	789	40	40	145	1	20	21	56
22	HTHBK35	209463 11/14/97	Uni-ZAP XR	32	884	1	884	108	108	146	1	26	27	66

Gene No.	cDNA Clone ID	ATCC Deposit Nr and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	First AA of Secreted Portion	Last AA of ORF
23	HWABA81	209463 11/14/97	pCMVSPORT 3.0	33	866	1	866	57	57	147	1	21	22	48
24	HKGAA73	209463 11/14/97	pSport1	34	1694	1	1694	38	38	148	1	27	28	88
25	HKIYP40	209463 11/14/97	pBluescript	35	1215	1	1215	43	43	149	1	32	33	76
26	HKMMW74	209463 11/14/97	pBluescript	36	1794	1	1794	202	202	150	1	21	22	41
27	HLFB127	209463 11/14/97	pBluescript SK-	37	1174	1	1174	135	135	151	1	19	20	45
28	HLQCW84	209463 11/14/97	Lambda ZAP II	38	1087	1	1087	31	31	152	1	18	19	41
29	HBNAV22	209463 11/14/97	Uni-ZAP XR	39	438	1	438	13	13	153	1	40	41	43
30	HTEAM34	209463 11/14/97	Uni-ZAP XR	40	734	1	734	63	63	154	1	28	29	122

Gene No.	cDNA Clone ID	ATCC Deposit Nr and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of 5' NT Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	First AA of Secreted Portion	Last AA of ORF
31	HTHDK34	209463 11/14/97	Uni-ZAP XR	41	1346	1	1346	60	60	155	1	35	36	41
32	H6BSG32	209463 11/14/97	Uni-ZAP XR	42	998	53	998	209	209	156	1	24	25	55
33	HCFAD33	209463 11/14/97	pSport1	43	658	1	658	297	297	157	1	17	18	44
34	HDTEN81	209463 11/14/97	pCMVSPORT 2.0	44	566	1	566	114	114	158	1	17	18	85
35	HFXDT43	209463 11/14/97	Lambda ZAP II	45	1277	1	1277	92	92	159	1	19	20	44
36	HNGHQ09	209463 11/14/97	Uni-ZAP XR	46	442	1	442	65	65	160	1	17	18	68
37	HHGDF16	209463 11/14/97	Lambda ZAP II	47	890	215	890	253	253	161	1	26	27	52
38	HJBCG12	209463 11/14/97	pBluescript SK-	48	737	40	737	382	382	162	1	24	25	56

Gene No.	cDNA Clone ID	ATCC Deposit Nr and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	First AA of Secreted Portion	Last AA of ORF
39	HOGAW62	209463 11/14/97	pCMVSPORT 2.0	49	571	1	571	259	259	163	1	26	27	55
40	HSWBJ74	209463 11/14/97	pCMVSPORT 3.0	50	356	1	356	43	43	164	1	35	36	47
41	HGBHR26	209511 12/03/97	Uni-ZAP XR	51	913	1	913	174	174	165	1	22	23	129
42	HKDBF34	209511 12/03/97	pCMVSPORT 1	52	1356	1	1356	18	18	166	1	19	20	104
43	H6EAB28	209511 12/03/97	Uni-ZAP XR	53	1547	1	1547	116	116	167	1	21	22	76
44	HLWAO22	209511 12/03/97	pCMVSPORT 3.0	54	1338	1	1311	212	212	168	1	21	22	354
45	HAGFH53	209511 12/03/97	Uni-ZAP XR	55	2071	1	2071	96	96	169	1	36	37	89
46	HHENQ22	209511 12/03/97	pCMVSPORT 3.0	56	1899	1	1899	115	115	170	1	36	37	58

Gene No.	eDNA Clone ID	ATCC Deposit Nr and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of 5' NT Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	First AA of Secreted Portion	Last AA of ORF
47	HKMLK53	209511 12/03/97	pBluescript	57	1543	1	1543	20	20	171	1	25	26	69
48	HSKGQ58	209511 12/03/97	pBluescript	58	1133	1	1133	41	41	172	1	37	38	78
49	HNFEQ93	209511 12/03/97	Uni-ZAP XR	59	1490	1	1480	33	33	173	1	33	34	173
50	HAIBZ39	209511 12/03/97	Uni-ZAP XR	60	1336	1	1336	48	48	174	1	17	18	63
51	HBXFP23	209511 12/03/97	ZAP Express	61	1705	178	1705	384	384	175	1	25	26	42
52	HEQBF32	209511 12/03/97	pCMVSPORT 3.0	62	1031	1	1031	97	97	176	1	25	26	113
53	HETHE81	209511 12/03/97	Uni-ZAP XR	63	1589	1	1589	182	182	177	1	23	24	155
54	HFPAC12	209511 12/03/97	Uni-ZAP XR	64	1088	1	1088	140	140	178	1	21	22	88

Gene No.	cDNA Clone ID	ATCC Deposit Nr and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	First AA of Secreted Portion	Last AA of ORF
55	HDPEF39	209511 12/03/97	pCMVSPORT 3.0	65	1256	1	1256	175	175	179	1	18	19	196
56	HFXHD88	209511 12/03/97	Lambda ZAP II	66	1602	1	1602	130	130	180	1	41	42	128
57	HFOXV65	209511 12/03/97	pSPORT1	67	938	1	938	204	204	181	1	26	27	154
58	HKADX21	209511 12/03/97	pCMVSPORT 2.0	68	1585	122	1585	414	414	182	1	18	19	106
59	HPZAB47	209511 12/03/97	pBluescript	69	1676	1	1676	34	34	183	1	18	19	47
60	HAGFE79	209511 12/03/97	Uni-ZAP XR	70	1344	1	1344	133	133	184	1	18	19	126
61	HCEIX60	209511 12/03/97	Uni-ZAP XR	71	1474	1	1474	38	38	185	1	25	26	86
62	HFXKD36	209511 12/03/97	Lambda ZAP II	72	2012	130	2012	251	251	186	1	35	36	57

Gene No.	cDNA Clone ID	ATCC Deposit Nr and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	First AA of Secreted Portion	Last AA of ORF
63	HBMCU71	209511 12/03/97	pBluescript	73	1267	1	1267	77	77	187	1	21	22	68
64	HTEIV80	209511 12/03/97	Uni-ZAP XR	74	1748	1	1748	203	203	188	1	15	16	47
65	HFIAP16	209511 12/03/97	pSport1	75	1570	38	1570	44	44	189	1	30	31	50
66	HODAV86	209511 12/03/97	Uni-ZAP XR	76	524	1	524	153	153	190	1	28	29	55
67	HTEDF80	209511 12/03/97	Uni-ZAP XR	77	1306	1	1306	696	696	191	1	21	22	126
68	HTODI69	209511 12/03/97	Uni-ZAP XR	78	1479	1	1479	123	123	192	1	23	24	69
69	HE6GR02	209511 12/03/97	Uni-ZAP XR	79	1794	1	1794	100	100	193	1	18	19	70
70	HAPNY86	209511 12/03/97	Uni-ZAP XR	80	1280	1	1280	100	100	194	1	25	26	129

Gene No.	cDNA Clone ID	ATCC Deposit Nr and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	First AA of Secreted Portion	Last AA of ORF
71	HTLDR33	209511 12/03/97	Uni-ZAP XR	81	974	1	974	368	368	195	1	19	20	54
72	HACB161	209511 12/03/97	Uni-ZAP XR	82	1955	1	1955	174	174	196	1	16	17	79
73	HMEIK34	209511 12/03/97	Lambda ZAP II	83	638	1	638	243	243	197	1	24	25	41
74	HKAAK02	209551 12/12/97	pCMVSPORT 2.0	84	859	1	859	97	97	198	1	33	34	196
75	HEPAA46	209551 12/12/97	Uni-ZAP XR	85	1129	1	1129	18	18	199	1	20	21	123
76	HFPCX09	209551 12/12/97	Uni-ZAP XR	86	2674	59	2674	249	249	200	1	26	27	549
76	HFPCX09	209551 12/12/97	Uni-ZAP XR	122	2207	1	2207	185	185	236	1	26	27	66
77	HLWAA88	209551 12/12/97	pCMVSPORT 3.0	87	1636	1	1636	51	51	201	1	22	23	488

Gene No.	cDNA Clone ID	ATCC Deposit Nr and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	First AA of Secreted Portion	Last AA of ORF
77	HLWAA88	209551 12/12/97	pCMVSPORT 3.0	123	1770	1	1770	35	35	237	1	22	23	113
78	HOHBV89	209551 12/12/97	pCMVSPORT 2.0	88	1639	1	1639	158	158	202	1	28	29	85
79	HHFBY69	209852 05/07/98	Uni-ZAP XR	89	1860	2	1860	71	71	203	1	25	26	399
79	HCEFL57	209551 12/12/97	Uni-ZAP XR	124	1034	1	1032	68	68	238	1	25	26	105
80	HMEKU83	209551 12/12/97	Lambda ZAP II	90	839	1	839	35	35	204	1	28	29	194
81	HOSBY40	209551 12/12/97	Uni-ZAP XR	91	1145	1	1145	89	89	205	1	30	31	56
82	HKFBH93	209551 12/12/97	ZAP Express	92	2050	174	2050	262	262	206	1	18	19	72
83	HMTAD67	209551 12/12/97	pCMVSPORT 3.0	93	1173	1	1173	306	306	207	1	19	20	84

Gene No.	cDNA Clone ID	ATCC Deposit Nr and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	First AA of Secreted Portion	Last AA of ORF
84	HTEBP77	209551 12/12/97	Uni-ZAP XR	94	822	1	822	91	91	208	1	18	19	194
85	HE9CO69	209551 12/12/97	Uni-ZAP XR	95	1077	1	1077	161	161	209	1	26	27	41
86	HCACV51	209551 12/12/97	Uni-ZAP XR	96	2092	1	2092	173	173	210	1	31	32	281
87	HPBP145	209551 12/12/97	Uni-ZAP XR	97	1352	1	1352	123	123	211	1	20	21	47
88	HLQDH79	209551 12/12/97	Lambda ZAP II	98	913	1	913	205	205	212	1	19	20	58
89	HNGFJ67	209551 12/12/97	Uni-ZAP XR	99	721	1	721	344	344	213	1	21	22	42
90	HEIAC52	209551 12/12/97	Uni-ZAP XR	100	645	1	645	81	81	214	1	18	19	52
91	HFXKL58	209551 12/12/97	Lambda ZAP II	101	563	1	563	120	120	215	1	17	18	67

Gene No.	cDNA Clone ID	ATCC Deposit Nr and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	5' NT of Clone Seq.	5' NT of Start Codon	5' NT of AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	First AA of Secreted Portion	Last AA of ORF
92	HMVAM60	209551 12/12/97	pSport1	102	1324	1	1324	96	96	216	1	26	27	56
93	HMVBR22	209551 12/12/97	pSport1	103	1731	1	1731	104	104	217	1	20	21	55
94	HPJCW04	209551 12/12/97	Uni-ZAP XR	104	1466	1	1466	44	44	218	1	19	20	57
95	HSIDJ81	209551 12/12/97	Uni-ZAP XR	105	1303	1	1303	8	8	219	1	22	23	58
96	HSLFU05	209551 12/12/97	Uni-ZAP XR	106	1516	1	1516	166	166	220	1	30	31	44
97	HEQAK71	209551 12/12/97	pCMVSPORT 3.0	107	1689	1	1689	198	198	221	1	17	18	44
98	HOSEQ49	209551 12/12/97	Uni-ZAP XR	108	1943	280	1935	544	544	222	1	32	33	51
99	HRAAM50	209551 12/12/97	pCMVSPORT 3.0	109	1594	1	1594	29	29	223	1	14	15	72

Gene No.	cDNA Clone ID	ATCC Deposit Nr and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of AA Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	First AA of Secreted Portion	Last AA of ORF
100	HSDFW45	209551 12/12/97	Uni-ZAP XR	110	1742	1	1742	118	118	224	1	19	20	70
101	HSLCQ82	209551 12/12/97	Uni-ZAP XR	111	1501	1	1501	233	233	225	1	22	23	57
102	HSSFT08	209551 12/12/97	Uni-ZAP XR	112	791	1	791	125	125	226	1	34	35	58
103	HTOIW31	209551 12/12/97	Uni-ZAP XR	113	1637	1	1637	107	107	227	1	19	20	42
104	HTXKQ85	209551 12/12/97	Uni-ZAP XR	114	1588	1	1588	61	61	228	1	28	29	40
105	HUFBK08	209551 12/12/97	pSport1	115	1926	1	1926	114	114	229	1	17	18	41
106	HAJAW31	209551 12/12/97	pCMVSPORT 3.0	116	1063	1	1063	41	41	230	1	22	23	164
107	HBUEE48	209551 12/12/97	Uni-ZAP XR	117	1615	1	1615	97	97	231	1	43	44	51

Gene No.	cDNA Clone ID	ATCC Deposit Nr and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT 3' NT of Clone Seq.	5' NT of Clone Seq.	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	First AA of Secreted Portion	Last AA of ORF
108	HBXGH74	209551 12/12/97	ZAP Express	118	1221	1	1221	188	232	1	20	21	129
109	HWBDM68	209551 12/12/97	pCMV Sport 3.0	119	1149	1	1130	368	233	1	24	25	54
110	HTPBW79	209511 12/03/97	Uni-ZAP XR	120	1515	118	1507	302	234	1	24	25	362

Table 1 summarizes the information corresponding to each "Gene No." described above. The nucleotide sequence identified as "NT SEQ ID NO:X" was assembled from partially homologous ("overlapping") sequences obtained from the "cDNA clone ID" identified in Table 1 and, in some cases, from additional related DNA clones. The
5 overlapping sequences were assembled into a single contiguous sequence of high redundancy (usually three to five overlapping sequences at each nucleotide position), resulting in a final sequence identified as SEQ ID NO:X.

The cDNA Clone ID was deposited on the date and given the corresponding deposit number listed in "ATCC Deposit No:Z and Date." Some of the deposits contain
10 multiple different clones corresponding to the same gene. "Vector" refers to the type of vector contained in the cDNA Clone ID.

"Total NT Seq." refers to the total number of nucleotides in the contig identified by "Gene No." The deposited clone may contain all or most of these sequences, reflected by the nucleotide position indicated as "5' NT of Clone Seq." and the "3' NT
15 of Clone Seq." of SEQ ID NO:X. The nucleotide position of SEQ ID NO:X of the putative start codon (methionine) is identified as "5' NT of Start Codon." Similarly, the nucleotide position of SEQ ID NO:X of the predicted signal sequence is identified as "5' NT of First AA of Signal Pep."

The translated amino acid sequence, beginning with the methionine, is identified
20 as "AA SEQ ID NO:Y," although other reading frames can also be easily translated using known molecular biology techniques. The polypeptides produced by these alternative open reading frames are specifically contemplated by the present invention.

The first and last amino acid position of SEQ ID NO:Y of the predicted signal peptide is identified as "First AA of Sig Pep" and "Last AA of Sig Pep." The predicted
25 first amino acid position of SEQ ID NO:Y of the secreted portion is identified as "Predicted First AA of Secreted Portion." Finally, the amino acid position of SEQ ID NO:Y of the last amino acid in the open reading frame is identified as "Last AA of ORF."

SEQ ID NO:X and the translated SEQ ID NO:Y are sufficiently accurate and
30 otherwise suitable for a variety of uses well known in the art and described further below. For instance, SEQ ID NO:X is useful for designing nucleic acid hybridization probes that will detect nucleic acid sequences contained in SEQ ID NO:X or the cDNA contained in the deposited clone. These probes will also hybridize to nucleic acid molecules in biological samples, thereby enabling a variety of forensic and diagnostic
35 methods of the invention. Similarly, polypeptides identified from SEQ ID NO:Y may be used to generate antibodies which bind specifically to the secreted proteins encoded by the cDNA clones identified in Table 1.

Nevertheless, DNA sequences generated by sequencing reactions can contain sequencing errors. The errors exist as misidentified nucleotides, or as insertions or deletions of nucleotides in the generated DNA sequence. The erroneously inserted or deleted nucleotides cause frame shifts in the reading frames of the predicted amino acid sequence. In these cases, the predicted amino acid sequence diverges from the actual amino acid sequence, even though the generated DNA sequence may be greater than 99.9% identical to the actual DNA sequence (for example, one base insertion or deletion in an open reading frame of over 1000 bases).

Accordingly, for those applications requiring precision in the nucleotide sequence or the amino acid sequence, the present invention provides not only the generated nucleotide sequence identified as SEQ ID NO:X and the predicted translated amino acid sequence identified as SEQ ID NO:Y, but also a sample of plasmid DNA containing a human cDNA of the invention deposited with the ATCC, as set forth in Table 1. The nucleotide sequence of each deposited clone can readily be determined by sequencing the deposited clone in accordance with known methods. The predicted amino acid sequence can then be verified from such deposits. Moreover, the amino acid sequence of the protein encoded by a particular clone can also be directly determined by peptide sequencing or by expressing the protein in a suitable host cell containing the deposited human cDNA, collecting the protein, and determining its sequence.

The present invention also relates to the genes corresponding to SEQ ID NO:X, SEQ ID NO:Y, or the deposited clone. The corresponding gene can be isolated in accordance with known methods using the sequence information disclosed herein. Such methods include preparing probes or primers from the disclosed sequence and identifying or amplifying the corresponding gene from appropriate sources of genomic material.

Also provided in the present invention are species homologs. Species homologs may be isolated and identified by making suitable probes or primers from the sequences provided herein and screening a suitable nucleic acid source for the desired homologue.

The polypeptides of the invention can be prepared in any suitable manner. Such polypeptides include isolated naturally occurring polypeptides, recombinantly produced polypeptides, synthetically produced polypeptides, or polypeptides produced by a combination of these methods. Means for preparing such polypeptides are well understood in the art.

The polypeptides may be in the form of the secreted protein, including the mature form, or may be a part of a larger protein, such as a fusion protein (see below).

It is often advantageous to include an additional amino acid sequence which contains secretory or leader sequences, pro-sequences, sequences which aid in purification, such as multiple histidine residues, or an additional sequence for stability during recombinant production.

- 5 The polypeptides of the present invention are preferably provided in an isolated form, and preferably are substantially purified. A recombinantly produced version of a polypeptide, including the secreted polypeptide, can be substantially purified by the one-step method described in Smith and Johnson, *Gene* 67:31-40 (1988). Polypeptides of the invention also can be purified from natural or recombinant sources using antibodies of the invention raised against the secreted protein in methods which are well known in the art.
- 10

Signal Sequences

- 15 Methods for predicting whether a protein has a signal sequence, as well as the cleavage point for that sequence, are available. For instance, the method of McGeoch, *Virus Res.* 3:271-286 (1985), uses the information from a short N-terminal charged region and a subsequent uncharged region of the complete (uncleaved) protein. The method of von Heinje, *Nucleic Acids Res.* 14:4683-4690 (1986) uses the information from the residues surrounding the cleavage site, typically residues -13 to +2, where +1 indicates the amino terminus of the secreted protein. The accuracy of predicting the cleavage points of known mammalian secretory proteins for each of these methods is in the range of 75-80%. (von Heinje, *supra*.) However, the two methods do not always produce the same predicted cleavage point(s) for a given protein.
- 20

- 25 In the present case, the deduced amino acid sequence of the secreted polypeptide was analyzed by a computer program called SignalP (Henrik Nielsen et al., *Protein Engineering* 10:1-6 (1997)), which predicts the cellular location of a protein based on the amino acid sequence. As part of this computational prediction of localization, the methods of McGeoch and von Heinje are incorporated. The analysis of the amino acid sequences of the secreted proteins described herein by this program provided the results shown in Table 1.
- 30

- As one of ordinary skill would appreciate, however, cleavage sites sometimes vary from organism to organism and cannot be predicted with absolute certainty. Accordingly, the present invention provides secreted polypeptides having a sequence shown in SEQ ID NO:Y which have an N-terminus beginning within 5 residues (i.e., + or - 5 residues) of the predicted cleavage point. Similarly, it is also recognized that in some cases, cleavage of the signal sequence from a secreted protein is not entirely
- 35

uniform, resulting in more than one secreted species. These polypeptides, and the polynucleotides encoding such polypeptides, are contemplated by the present invention.

Moreover, the signal sequence identified by the above analysis may not necessarily predict the naturally occurring signal sequence. For example, the naturally occurring signal sequence may be further upstream from the predicted signal sequence. However, it is likely that the predicted signal sequence will be capable of directing the secreted protein to the ER. These polypeptides, and the polynucleotides encoding such polypeptides, are contemplated by the present invention.

10 Polynucleotide and Polypeptide Variants

"Variant" refers to a polynucleotide or polypeptide differing from the polynucleotide or polypeptide of the present invention, but retaining essential properties thereof. Generally, variants are overall closely similar, and, in many regions, identical to the polynucleotide or polypeptide of the present invention.

15 By a polynucleotide having a nucleotide sequence at least, for example, 95% "identical" to a reference nucleotide sequence of the present invention, it is intended that the nucleotide sequence of the polynucleotide is identical to the reference sequence except that the polynucleotide sequence may include up to five point mutations per each 100 nucleotides of the reference nucleotide sequence encoding the polypeptide. In other words, to obtain a polynucleotide having a nucleotide sequence at least 95% identical to a reference nucleotide sequence, up to 5% of the nucleotides in the reference sequence may be deleted or substituted with another nucleotide, or a number of nucleotides up to 5% of the total nucleotides in the reference sequence may be inserted into the reference sequence. The query sequence may be an entire sequence shown in Table 1, the ORF (open reading frame), or any fragment specified as described herein.

25 As a practical matter, whether any particular nucleic acid molecule or polypeptide is at least 90%, 95%, 96%, 97%, 98% or 99% identical to a nucleotide sequence of the present invention can be determined conventionally using known computer programs. A preferred method for determining the best overall match between a query sequence (a sequence of the present invention) and a subject sequence, also referred to as a global sequence alignment, can be determined using the FASTDB computer program based on the algorithm of Brutlag et al. (Comp. App. Biosci. (1990) 6:237-245). In a sequence alignment the query and subject sequences are both DNA sequences. An RNA sequence can be compared by converting U's to T's. The result of said global sequence alignment is in percent identity. Preferred parameters used in a FASTDB alignment of DNA sequences to calculate percent identity are:

35 Matrix=Unitary, k-tuple=4, Mismatch Penalty=1, Joining Penalty=30, Randomization

Group Length=0, Cutoff Score=1, Gap Penalty=5, Gap Size Penalty 0.05, Window Size=500 or the length of the subject nucleotide sequence, whichever is shorter.

If the subject sequence is shorter than the query sequence because of 5' or 3' deletions, not because of internal deletions, a manual correction must be made to the results. This is because the FASTDB program does not account for 5' and 3' truncations of the subject sequence when calculating percent identity. For subject sequences truncated at the 5' or 3' ends, relative to the query sequence, the percent identity is corrected by calculating the number of bases of the query sequence that are 5' and 3' of the subject sequence, which are not matched/aligned, as a percent of the total bases of the query sequence. Whether a nucleotide is matched/aligned is determined by results of the FASTDB sequence alignment. This percentage is then subtracted from the percent identity, calculated by the above FASTDB program using the specified parameters, to arrive at a final percent identity score. This corrected score is what is used for the purposes of the present invention. Only bases outside the 5' and 3' bases of the subject sequence, as displayed by the FASTDB alignment, which are not matched/aligned with the query sequence, are calculated for the purposes of manually adjusting the percent identity score.

For example, a 90 base subject sequence is aligned to a 100 base query sequence to determine percent identity. The deletions occur at the 5' end of the subject sequence and therefore, the FASTDB alignment does not show a matched/alignment of the first 10 bases at 5' end. The 10 unpaired bases represent 10% of the sequence (number of bases at the 5' and 3' ends not matched/total number of bases in the query sequence) so 10% is subtracted from the percent identity score calculated by the FASTDB program. If the remaining 90 bases were perfectly matched the final percent identity would be 90%. In another example, a 90 base subject sequence is compared with a 100 base query sequence. This time the deletions are internal deletions so that there are no bases on the 5' or 3' of the subject sequence which are not matched/aligned with the query. In this case the percent identity calculated by FASTDB is not manually corrected. Once again, only bases 5' and 3' of the subject sequence which are not matched/aligned with the query sequence are manually corrected for. No other manual corrections are to be made for the purposes of the present invention.

By a polypeptide having an amino acid sequence at least, for example, 95% "identical" to a query amino acid sequence of the present invention, it is intended that the amino acid sequence of the subject polypeptide is identical to the query sequence except that the subject polypeptide sequence may include up to five amino acid alterations per each 100 amino acids of the query amino acid sequence. In other words, to obtain a polypeptide having an amino acid sequence at least 95% identical to a query

amino acid sequence, up to 5% of the amino acid residues in the subject sequence may be inserted, deleted, (indels) or substituted with another amino acid. These alterations of the reference sequence may occur at the amino or carboxy terminal positions of the reference amino acid sequence or anywhere between those terminal positions, interspersed either individually among residues in the reference sequence or in one or more contiguous groups within the reference sequence.

As a practical matter, whether any particular polypeptide is at least 90%, 95%, 96%, 97%, 98% or 99% identical to, for instance, the amino acid sequences shown in Table 1 or to the amino acid sequence encoded by deposited DNA clone can be determined conventionally using known computer programs. A preferred method for determining the best overall match between a query sequence (a sequence of the present invention) and a subject sequence, also referred to as a global sequence alignment, can be determined using the FASTDB computer program based on the algorithm of Brutlag et al. (Comp. App. Biosci. (1990) 6:237-245). In a sequence alignment the query and subject sequences are either both nucleotide sequences or both amino acid sequences. The result of said global sequence alignment is in percent identity. Preferred parameters used in a FASTDB amino acid alignment are: Matrix=PAM 0, k-tuple=2, Mismatch Penalty=1, Joining Penalty=20, Randomization Group Length=0, Cutoff Score=1, Window Size=sequence length, Gap Penalty=5, Gap Size Penalty=0.05, Window Size=500 or the length of the subject amino acid sequence, whichever is shorter.

If the subject sequence is shorter than the query sequence due to N- or C-terminal deletions, not because of internal deletions, a manual correction must be made to the results. This is because the FASTDB program does not account for N- and C-terminal truncations of the subject sequence when calculating global percent identity. For subject sequences truncated at the N- and C-termini, relative to the query sequence, the percent identity is corrected by calculating the number of residues of the query sequence that are N- and C-terminal of the subject sequence, which are not matched/aligned with a corresponding subject residue, as a percent of the total bases of the query sequence. Whether a residue is matched/aligned is determined by results of the FASTDB sequence alignment. This percentage is then subtracted from the percent identity, calculated by the above FASTDB program using the specified parameters, to arrive at a final percent identity score. This final percent identity score is what is used for the purposes of the present invention. Only residues to the N- and C-termini of the subject sequence, which are not matched/aligned with the query sequence, are considered for the purposes of manually adjusting the percent identity score. That is, only query residue positions outside the farthest N- and C-terminal residues of the subject sequence.

For example, a 90 amino acid residue subject sequence is aligned with a 100 residue query sequence to determine percent identity. The deletion occurs at the N-terminus of the subject sequence and therefore, the FASTDB alignment does not show a matching/alignment of the first 10 residues at the N-terminus. The 10 unpaired
5 residues represent 10% of the sequence (number of residues at the N- and C- termini not matched/total number of residues in the query sequence) so 10% is subtracted from the percent identity score calculated by the FASTDB program. If the remaining 90 residues were perfectly matched the final percent identity would be 90%. In another example, a 90 residue subject sequence is compared with a 100 residue query sequence.
10 This time the deletions are internal deletions so there are no residues at the N- or C-termini of the subject sequence which are not matched/aligned with the query. In this case the percent identity calculated by FASTDB is not manually corrected. Once again, only residue positions outside the N- and C-terminal ends of the subject sequence, as displayed in the FASTDB alignment, which are not matched/aligned with the query
15 sequence are manually corrected for. No other manual corrections are to be made for the purposes of the present invention.

The variants may contain alterations in the coding regions, non-coding regions, or both. Especially preferred are polynucleotide variants containing alterations which produce silent substitutions, additions, or deletions, but do not alter the properties or
20 activities of the encoded polypeptide. Nucleotide variants produced by silent substitutions due to the degeneracy of the genetic code are preferred. Moreover, variants in which 5-10, 1-5, or 1-2 amino acids are substituted, deleted, or added in any combination are also preferred. Polynucleotide variants can be produced for a variety of reasons, e.g., to optimize codon expression for a particular host (change codons in
25 the human mRNA to those preferred by a bacterial host such as *E. coli*).

Naturally occurring variants are called "allelic variants," and refer to one of several alternate forms of a gene occupying a given locus on a chromosome of an organism. (Genes II, Lewin, B., ed., John Wiley & Sons, New York (1985).) These allelic variants can vary at either the polynucleotide and/or polypeptide level.
30 Alternatively, non-naturally occurring variants may be produced by mutagenesis techniques or by direct synthesis.

Using known methods of protein engineering and recombinant DNA technology, variants may be generated to improve or alter the characteristics of the polypeptides of the present invention. For instance, one or more amino acids can be
35 deleted from the N-terminus or C-terminus of the secreted protein without substantial loss of biological function. The authors of Ron et al., *J. Biol. Chem.* 268: 2984-2988 (1993), reported variant KGF proteins having heparin binding activity even after

deleting 3, 8, or 27 amino-terminal amino acid residues. Similarly, Interferon gamma exhibited up to ten times higher activity after deleting 8-10 amino acid residues from the carboxy terminus of this protein. (Dobeli et al., J. Biotechnology 7:199-216 (1988).)

Moreover, ample evidence demonstrates that variants often retain a biological activity similar to that of the naturally occurring protein. For example, Gayle and coworkers (J. Biol. Chem 268:22105-22111 (1993)) conducted extensive mutational analysis of human cytokine IL-1a. They used random mutagenesis to generate over 3,500 individual IL-1a mutants that averaged 2.5 amino acid changes per variant over the entire length of the molecule. Multiple mutations were examined at every possible amino acid position. The investigators found that "[m]ost of the molecule could be altered with little effect on either [binding or biological activity]." (See, Abstract.) In fact, only 23 unique amino acid sequences, out of more than 3,500 nucleotide sequences examined, produced a protein that significantly differed in activity from wild-type.

Furthermore, even if deleting one or more amino acids from the N-terminus or C-terminus of a polypeptide results in modification or loss of one or more biological functions, other biological activities may still be retained. For example, the ability of a deletion variant to induce and/or to bind antibodies which recognize the secreted form will likely be retained when less than the majority of the residues of the secreted form are removed from the N-terminus or C-terminus. Whether a particular polypeptide lacking N- or C-terminal residues of a protein retains such immunogenic activities can readily be determined by routine methods described herein and otherwise known in the art.

Thus, the invention further includes polypeptide variants which show substantial biological activity. Such variants include deletions, insertions, inversions, repeats, and substitutions selected according to general rules known in the art so as have little effect on activity. For example, guidance concerning how to make phenotypically silent amino acid substitutions is provided in Bowie, J. U. et al., Science 247:1306-1310 (1990), wherein the authors indicate that there are two main strategies for studying the tolerance of an amino acid sequence to change.

The first strategy exploits the tolerance of amino acid substitutions by natural selection during the process of evolution. By comparing amino acid sequences in different species, conserved amino acids can be identified. These conserved amino acids are likely important for protein function. In contrast, the amino acid positions where substitutions have been tolerated by natural selection indicates that these positions are not critical for protein function. Thus, positions tolerating amino acid substitution could be modified while still maintaining biological activity of the protein.

The second strategy uses genetic engineering to introduce amino acid changes at specific positions of a cloned gene to identify regions critical for protein function. For example, site directed mutagenesis or alanine-scanning mutagenesis (introduction of single alanine mutations at every residue in the molecule) can be used. (Cunningham and Wells, Science 244:1081-1085 (1989).) The resulting mutant molecules can then be tested for biological activity.

As the authors state, these two strategies have revealed that proteins are surprisingly tolerant of amino acid substitutions. The authors further indicate which amino acid changes are likely to be permissive at certain amino acid positions in the protein. For example, most buried (within the tertiary structure of the protein) amino acid residues require nonpolar side chains, whereas few features of surface side chains are generally conserved. Moreover, tolerated conservative amino acid substitutions involve replacement of the aliphatic or hydrophobic amino acids Ala, Val, Leu and Ile; replacement of the hydroxyl residues Ser and Thr; replacement of the acidic residues Asp and Glu; replacement of the amide residues Asn and Gln, replacement of the basic residues Lys, Arg, and His; replacement of the aromatic residues Phe, Tyr, and Trp, and replacement of the small-sized amino acids Ala, Ser, Thr, Met, and Gly.

Besides conservative amino acid substitution, variants of the present invention include (i) substitutions with one or more of the non-conserved amino acid residues, where the substituted amino acid residues may or may not be one encoded by the genetic code, or (ii) substitution with one or more of amino acid residues having a substituent group, or (iii) fusion of the mature polypeptide with another compound, such as a compound to increase the stability and/or solubility of the polypeptide (for example, polyethylene glycol), or (iv) fusion of the polypeptide with additional amino acids, such as an IgG Fc fusion region peptide, or leader or secretory sequence, or a sequence facilitating purification. Such variant polypeptides are deemed to be within the scope of those skilled in the art from the teachings herein.

For example, polypeptide variants containing amino acid substitutions of charged amino acids with other charged or neutral amino acids may produce proteins with improved characteristics, such as less aggregation. Aggregation of pharmaceutical formulations both reduces activity and increases clearance due to the aggregate's immunogenic activity. (Pinckard et al., Clin. Exp. Immunol. 2:331-340 (1967); Robbins et al., Diabetes 36: 838-845 (1987); Cleland et al., Crit. Rev. Therapeutic Drug Carrier Systems 10:307-377 (1993).)

Polynucleotide and Polypeptide Fragments

In the present invention, a "polynucleotide fragment" refers to a short polynucleotide having a nucleic acid sequence contained in the deposited clone or shown in SEQ ID NO:X. The short nucleotide fragments are preferably at least about 15 nt, and more preferably at least about 20 nt, still more preferably at least about 30 nt, and even more preferably, at least about 40 nt in length. A fragment "at least 20 nt in length," for example, is intended to include 20 or more contiguous bases from the cDNA sequence contained in the deposited clone or the nucleotide sequence shown in SEQ ID NO:X. These nucleotide fragments are useful as diagnostic probes and primers as discussed herein. Of course, larger fragments (e.g., 50, 150, 500, 600, 2000 nucleotides) are preferred.

Moreover, representative examples of polynucleotide fragments of the invention, include, for example, fragments having a sequence from about nucleotide number 1-50, 51-100, 101-150, 151-200, 201-250, 251-300, 301-350, 351-400, 401-450, 451-500, 501-550, 551-600, 651-700, 701-750, 751-800, 800-850, 851-900, 901-950, 951-1000, 1001-1050, 1051-1100, 1101-1150, 1151-1200, 1201-1250, 1251-1300, 1301-1350, 1351-1400, 1401-1450, 1451-1500, 1501-1550, 1551-1600, 1601-1650, 1651-1700, 1701-1750, 1751-1800, 1801-1850, 1851-1900, 1901-1950, 1951-2000, or 2001 to the end of SEQ ID NO:X or the cDNA contained in the deposited clone. In this context "about" includes the particularly recited ranges, larger or smaller by several (5, 4, 3, 2, or 1) nucleotides, at either terminus or at both termini. Preferably, these fragments encode a polypeptide which has biological activity. More preferably, these polynucleotides can be used as probes or primers as discussed herein.

In the present invention, a "polypeptide fragment" refers to a short amino acid sequence contained in SEQ ID NO:Y or encoded by the cDNA contained in the deposited clone. Protein fragments may be "free-standing," or comprised within a larger polypeptide of which the fragment forms a part or region, most preferably as a single continuous region. Representative examples of polypeptide fragments of the invention, include, for example, fragments from about amino acid number 1-20, 21-40, 41-60, 61-80, 81-100, 102-120, 121-140, 141-160, or 161 to the end of the coding region. Moreover, polypeptide fragments can be about 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, or 150 amino acids in length. In this context "about" includes the particularly recited ranges, larger or smaller by several (5, 4, 3, 2, or 1) amino acids, at either extreme or at both extremes.

Preferred polypeptide fragments include the secreted protein as well as the mature form. Further preferred polypeptide fragments include the secreted protein or the mature form having a continuous series of deleted residues from the amino or the carboxy terminus, or both. For example, any number of amino acids, ranging from 1-

60, can be deleted from the amino terminus of either the secreted polypeptide or the mature form. Similarly, any number of amino acids, ranging from 1-30, can be deleted from the carboxy terminus of the secreted protein or mature form. Furthermore, any combination of the above amino and carboxy terminus deletions are preferred.

- 5 Similarly, polynucleotide fragments encoding these polypeptide fragments are also preferred.

Also preferred are polypeptide and polynucleotide fragments characterized by structural or functional domains, such as fragments that comprise alpha-helix and alpha-helix forming regions, beta-sheet and beta-sheet-forming regions, turn and turn-
10 forming regions, coil and coil-forming regions, hydrophilic regions, hydrophobic regions, alpha amphipathic regions, beta amphipathic regions, flexible regions, surface-forming regions, substrate binding region, and high antigenic index regions. Polypeptide fragments of SEQ ID NO:Y falling within conserved domains are specifically contemplated by the present invention. Moreover, polynucleotide
15 fragments encoding these domains are also contemplated.

Other preferred fragments are biologically active fragments. Biologically active fragments are those exhibiting activity similar, but not necessarily identical, to an activity of the polypeptide of the present invention. The biological activity of the fragments may include an improved desired activity, or a decreased undesirable activity.
20

Epitopes & Antibodies

In the present invention, "epitopes" refer to polypeptide fragments having antigenic or immunogenic activity in an animal, especially in a human. A preferred embodiment of the present invention relates to a polypeptide fragment comprising an
25 epitope, as well as the polynucleotide encoding this fragment. A region of a protein molecule to which an antibody can bind is defined as an "antigenic epitope." In contrast, an "immunogenic epitope" is defined as a part of a protein that elicits an antibody response. (See, for instance, Geysen et al., Proc. Natl. Acad. Sci. USA 81:3998- 4002 (1983).)

30 Fragments which function as epitopes may be produced by any conventional means. (See, e.g., Houghten, R. A., Proc. Natl. Acad. Sci. USA 82:5131-5135 (1985) further described in U.S. Patent No. 4,631,211.)

In the present invention, antigenic epitopes preferably contain a sequence of at least seven, more preferably at least nine, and most preferably between about 15 to
35 about 30 amino acids. Antigenic epitopes are useful to raise antibodies, including monoclonal antibodies, that specifically bind the epitope. (See, for instance, Wilson et al., Cell 37:767-778 (1984); Sutcliffe, J. G. et al., Science 219:660-666 (1983).)

Similarly, immunogenic epitopes can be used to induce antibodies according to methods well known in the art. (See, for instance, Sutcliffe et al., supra; Wilson et al., supra; Chow, M. et al., Proc. Natl. Acad. Sci. USA 82:910-914; and Bittle, F. J. et al., J. Gen. Virol. 66:2347-2354 (1985).) A preferred immunogenic epitope includes the secreted protein. The immunogenic epitopes may be presented together with a carrier protein, such as an albumin, to an animal system (such as rabbit or mouse) or, if it is long enough (at least about 25 amino acids), without a carrier. However, immunogenic epitopes comprising as few as 8 to 10 amino acids have been shown to be sufficient to raise antibodies capable of binding to, at the very least, linear epitopes in a denatured polypeptide (e.g., in Western blotting.)

As used herein, the term "antibody" (Ab) or "monoclonal antibody" (Mab) is meant to include intact molecules as well as antibody fragments (such as, for example, Fab and F(ab')₂ fragments) which are capable of specifically binding to protein. Fab and F(ab')₂ fragments lack the Fc fragment of intact antibody, clear more rapidly from the circulation, and may have less non-specific tissue binding than an intact antibody. (Wahl et al., J. Nucl. Med. 24:316-325 (1983).) Thus, these fragments are preferred, as well as the products of a FAB or other immunoglobulin expression library. Moreover, antibodies of the present invention include chimeric, single chain, and humanized antibodies.

20

Fusion Proteins

Any polypeptide of the present invention can be used to generate fusion proteins. For example, the polypeptide of the present invention, when fused to a second protein, can be used as an antigenic tag. Antibodies raised against the polypeptide of the present invention can be used to indirectly detect the second protein by binding to the polypeptide. Moreover, because secreted proteins target cellular locations based on trafficking signals, the polypeptides of the present invention can be used as targeting molecules once fused to other proteins.

Examples of domains that can be fused to polypeptides of the present invention include not only heterologous signal sequences, but also other heterologous functional regions. The fusion does not necessarily need to be direct, but may occur through linker sequences.

Moreover, fusion proteins may also be engineered to improve characteristics of the polypeptide of the present invention. For instance, a region of additional amino acids, particularly charged amino acids, may be added to the N-terminus of the polypeptide to improve stability and persistence during purification from the host cell or subsequent handling and storage. Also, peptide moieties may be added to the

polypeptide to facilitate purification. Such regions may be removed prior to final preparation of the polypeptide. The addition of peptide moieties to facilitate handling of polypeptides are familiar and routine techniques in the art.

Moreover, polypeptides of the present invention, including fragments, and specifically epitopes, can be combined with parts of the constant domain of immunoglobulins (IgG), resulting in chimeric polypeptides. These fusion proteins facilitate purification and show an increased half-life in vivo. One reported example describes chimeric proteins consisting of the first two domains of the human CD4-polypeptide and various domains of the constant regions of the heavy or light chains of mammalian immunoglobulins. (EP A 394,827; Traunecker et al., *Nature* 331:84-86 (1988).) Fusion proteins having disulfide-linked dimeric structures (due to the IgG) can also be more efficient in binding and neutralizing other molecules, than the monomeric secreted protein or protein fragment alone. (Fountoulakis et al., *J. Biochem.* 270:3958-3964 (1995).)

Similarly, EP-A-O 464 533 (Canadian counterpart 2045869) discloses fusion proteins comprising various portions of constant region of immunoglobulin molecules together with another human protein or part thereof. In many cases, the Fc part in a fusion protein is beneficial in therapy and diagnosis, and thus can result in, for example, improved pharmacokinetic properties. (EP-A 0232 262.) Alternatively, deleting the Fc part after the fusion protein has been expressed, detected, and purified, would be desired. For example, the Fc portion may hinder therapy and diagnosis if the fusion protein is used as an antigen for immunizations. In drug discovery, for example, human proteins, such as hIL-5, have been fused with Fc portions for the purpose of high-throughput screening assays to identify antagonists of hIL-5. (See, D. Bennett et al., *J. Molecular Recognition* 8:52-58 (1995); K. Johanson et al., *J. Biol. Chem.* 270:9459-9471 (1995).)

Moreover, the polypeptides of the present invention can be fused to marker sequences, such as a peptide which facilitates purification of the fused polypeptide. In preferred embodiments, the marker amino acid sequence is a hexa-histidine peptide, such as the tag provided in a pQE vector (QIAGEN, Inc., 9259 Eton Avenue, Chatsworth, CA, 91311), among others, many of which are commercially available. As described in Gentz et al., *Proc. Natl. Acad. Sci. USA* 86:821-824 (1989), for instance, hexa-histidine provides for convenient purification of the fusion protein. Another peptide tag useful for purification, the "HA" tag, corresponds to an epitope derived from the influenza hemagglutinin protein. (Wilson et al., *Cell* 37:767 (1984).)

Thus, any of these above fusions can be engineered using the polynucleotides or the polypeptides of the present invention.

Vectors, Host Cells, and Protein Production

The present invention also relates to vectors containing the polynucleotide of the present invention, host cells, and the production of polypeptides by recombinant techniques. The vector may be, for example, a phage, plasmid, viral, or retroviral vector. Retroviral vectors may be replication competent or replication defective. In the latter case, viral propagation generally will occur only in complementing host cells.

The polynucleotides may be joined to a vector containing a selectable marker for propagation in a host. Generally, a plasmid vector is introduced in a precipitate, such as a calcium phosphate precipitate, or in a complex with a charged lipid. If the vector is a virus, it may be packaged in vitro using an appropriate packaging cell line and then transduced into host cells.

The polynucleotide insert should be operatively linked to an appropriate promoter, such as the phage lambda PL promoter, the E. coli lac, trp, phoA and tac promoters, the SV40 early and late promoters and promoters of retroviral LTRs, to name a few. Other suitable promoters will be known to the skilled artisan. The expression constructs will further contain sites for transcription initiation, termination, and, in the transcribed region, a ribosome binding site for translation. The coding portion of the transcripts expressed by the constructs will preferably include a translation initiating codon at the beginning and a termination codon (UAA, UGA or UAG) appropriately positioned at the end of the polypeptide to be translated.

As indicated, the expression vectors will preferably include at least one selectable marker. Such markers include dihydrofolate reductase, G418 or neomycin resistance for eukaryotic cell culture and tetracycline, kanamycin or ampicillin resistance genes for culturing in E. coli and other bacteria. Representative examples of appropriate hosts include, but are not limited to, bacterial cells, such as E. coli, Streptomyces and Salmonella typhimurium cells; fungal cells, such as yeast cells; insect cells such as Drosophila S2 and Spodoptera Sf9 cells; animal cells such as CHO, COS, 293, and Bowes melanoma cells; and plant cells. Appropriate culture mediums and conditions for the above-described host cells are known in the art.

Among vectors preferred for use in bacteria include pQE70, pQE60 and pQE-9, available from QIAGEN, Inc.; pBluescript vectors, Phagescript vectors, pNH8A, pNH16a, pNH18A, pNH46A, available from Stratagene Cloning Systems, Inc.; and ptrc99a, pKK223-3, pKK233-3, pDR540, pRIT5 available from Pharmacia Biotech, Inc. Among preferred eukaryotic vectors are pWLNEO, pSV2CAT, pOG44, pXT1 and pSG available from Stratagene; and pSVK3, pBPV, pMSG and pSVL available from Pharmacia. Other suitable vectors will be readily apparent to the skilled artisan.

Introduction of the construct into the host cell can be effected by calcium phosphate transfection, DEAE-dextran mediated transfection, cationic lipid-mediated transfection, electroporation, transduction, infection, or other methods. Such methods are described in many standard laboratory manuals, such as Davis et al., Basic Methods
5 In Molecular Biology (1986). It is specifically contemplated that the polypeptides of the present invention may in fact be expressed by a host cell lacking a recombinant vector.

A polypeptide of this invention can be recovered and purified from recombinant cell cultures by well-known methods including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography,
10 phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography and lectin chromatography. Most preferably, high performance liquid chromatography ("HPLC") is employed for purification.

Polypeptides of the present invention, and preferably the secreted form, can also
15 be recovered from: products purified from natural sources, including bodily fluids, tissues and cells, whether directly isolated or cultured; products of chemical synthetic procedures; and products produced by recombinant techniques from a prokaryotic or eukaryotic host, including, for example, bacterial, yeast, higher plant, insect, and mammalian cells. Depending upon the host employed in a recombinant production
20 procedure, the polypeptides of the present invention may be glycosylated or may be non-glycosylated. In addition, polypeptides of the invention may also include an initial modified methionine residue, in some cases as a result of host-mediated processes. Thus, it is well known in the art that the N-terminal methionine encoded by the translation initiation codon generally is removed with high efficiency from any protein
25 after translation in all eukaryotic cells. While the N-terminal methionine on most proteins also is efficiently removed in most prokaryotes, for some proteins, this prokaryotic removal process is inefficient, depending on the nature of the amino acid to which the N-terminal methionine is covalently linked.

30 Uses of the Polynucleotides

Each of the polynucleotides identified herein can be used in numerous ways as reagents. The following description should be considered exemplary and utilizes known techniques.

The polynucleotides of the present invention are useful for chromosome
35 identification. There exists an ongoing need to identify new chromosome markers, since few chromosome marking reagents, based on actual sequence data (repeat

polymorphisms), are presently available. Each polynucleotide of the present invention can be used as a chromosome marker.

Briefly, sequences can be mapped to chromosomes by preparing PCR primers (preferably 15-25 bp) from the sequences shown in SEQ ID NO:X. Primers can be
5 selected using computer analysis so that primers do not span more than one predicted exon in the genomic DNA. These primers are then used for PCR screening of somatic cell hybrids containing individual human chromosomes. Only those hybrids containing the human gene corresponding to the SEQ ID NO:X will yield an amplified fragment.

Similarly, somatic hybrids provide a rapid method of PCR mapping the
10 polynucleotides to particular chromosomes. Three or more clones can be assigned per day using a single thermal cycler. Moreover, sublocalization of the polynucleotides can be achieved with panels of specific chromosome fragments. Other gene mapping strategies that can be used include in situ hybridization, prescreening with labeled flow-sorted chromosomes, and preselection by hybridization to construct chromosome
15 specific-cDNA libraries.

Precise chromosomal location of the polynucleotides can also be achieved using fluorescence in situ hybridization (FISH) of a metaphase chromosomal spread. This technique uses polynucleotides as short as 500 or 600 bases; however, polynucleotides 2,000-4,000 bp are preferred. For a review of this technique, see Verma et al.,
20 "Human Chromosomes: a Manual of Basic Techniques," Pergamon Press, New York (1988).

For chromosome mapping, the polynucleotides can be used individually (to mark a single chromosome or a single site on that chromosome) or in panels (for marking multiple sites and/or multiple chromosomes). Preferred polynucleotides
25 correspond to the noncoding regions of the cDNAs because the coding sequences are more likely conserved within gene families, thus increasing the chance of cross hybridization during chromosomal mapping.

Once a polynucleotide has been mapped to a precise chromosomal location, the physical position of the polynucleotide can be used in linkage analysis. Linkage
30 analysis establishes coinheritance between a chromosomal location and presentation of a particular disease. (Disease mapping data are found, for example, in V. McKusick, Mendelian Inheritance in Man (available on line through Johns Hopkins University Welch Medical Library) .) Assuming 1 megabase mapping resolution and one gene per 20 kb, a cDNA precisely localized to a chromosomal region associated with the disease
35 could be one of 50-500 potential causative genes.

Thus, once coinheritance is established, differences in the polynucleotide and the corresponding gene between affected and unaffected individuals can be examined.

First, visible structural alterations in the chromosomes, such as deletions or translocations, are examined in chromosome spreads or by PCR. If no structural alterations exist, the presence of point mutations are ascertained. Mutations observed in some or all affected individuals, but not in normal individuals, indicates that the
5 mutation may cause the disease. However, complete sequencing of the polypeptide and the corresponding gene from several normal individuals is required to distinguish the mutation from a polymorphism. If a new polymorphism is identified, this polymorphic polypeptide can be used for further linkage analysis.

Furthermore, increased or decreased expression of the gene in affected
10 individuals as compared to unaffected individuals can be assessed using polynucleotides of the present invention. Any of these alterations (altered expression, chromosomal rearrangement, or mutation) can be used as a diagnostic or prognostic marker.

In addition to the foregoing, a polynucleotide can be used to control gene
15 expression through triple helix formation or antisense DNA or RNA. Both methods rely on binding of the polynucleotide to DNA or RNA. For these techniques, preferred polynucleotides are usually 20 to 40 bases in length and complementary to either the region of the gene involved in transcription (triple helix - see Lee et al., Nucl. Acids Res. 6:3073 (1979); Cooney et al., Science 241:456 (1988); and Dervan et al., Science
20 251:1360 (1991)) or to the mRNA itself (antisense - Okano, J. Neurochem. 56:560 (1991); Oligodeoxy-nucleotides as Antisense Inhibitors of Gene Expression, CRC Press, Boca Raton, FL (1988).) Triple helix formation optimally results in a shut-off of RNA transcription from DNA, while antisense RNA hybridization blocks translation of an mRNA molecule into polypeptide. Both techniques are effective in model
25 systems, and the information disclosed herein can be used to design antisense or triple helix polynucleotides in an effort to treat disease.

Polynucleotides of the present invention are also useful in gene therapy. One goal of gene therapy is to insert a normal gene into an organism having a defective gene, in an effort to correct the genetic defect. The polynucleotides disclosed in the
30 present invention offer a means of targeting such genetic defects in a highly accurate manner. Another goal is to insert a new gene that was not present in the host genome, thereby producing a new trait in the host cell.

The polynucleotides are also useful for identifying individuals from minute biological samples. The United States military, for example, is considering the use of
35 restriction fragment length polymorphism (RFLP) for identification of its personnel. In this technique, an individual's genomic DNA is digested with one or more restriction enzymes, and probed on a Southern blot to yield unique bands for identifying

personnel. This method does not suffer from the current limitations of "Dog Tags" which can be lost, switched, or stolen, making positive identification difficult. The polynucleotides of the present invention can be used as additional DNA markers for RFLP.

5 The polynucleotides of the present invention can also be used as an alternative to RFLP, by determining the actual base-by-base DNA sequence of selected portions of an individual's genome. These sequences can be used to prepare PCR primers for amplifying and isolating such selected DNA, which can then be sequenced. Using this technique, individuals can be identified because each individual will have a unique set
10 of DNA sequences. Once an unique ID database is established for an individual, positive identification of that individual, living or dead, can be made from extremely small tissue samples.

Forensic biology also benefits from using DNA-based identification techniques as disclosed herein. DNA sequences taken from very small biological samples such as
15 tissues, e.g., hair or skin, or body fluids, e.g., blood, saliva, semen, etc., can be amplified using PCR. In one prior art technique, gene sequences amplified from polymorphic loci, such as DQa class II HLA gene, are used in forensic biology to identify individuals. (Erlich, H., PCR Technology, Freeman and Co. (1992).) Once these specific polymorphic loci are amplified, they are digested with one or more
20 restriction enzymes, yielding an identifying set of bands on a Southern blot probed with DNA corresponding to the DQa class II HLA gene. Similarly, polynucleotides of the present invention can be used as polymorphic markers for forensic purposes.

There is also a need for reagents capable of identifying the source of a particular tissue. Such need arises, for example, in forensics when presented with tissue of
25 unknown origin. Appropriate reagents can comprise, for example, DNA probes or primers specific to particular tissue prepared from the sequences of the present invention. Panels of such reagents can identify tissue by species and/or by organ type. In a similar fashion, these reagents can be used to screen tissue cultures for contamination.

30 In the very least, the polynucleotides of the present invention can be used as molecular weight markers on Southern gels, as diagnostic probes for the presence of a specific mRNA in a particular cell type, as a probe to "subtract-out" known sequences in the process of discovering novel polynucleotides, for selecting and making oligomers for attachment to a "gene chip" or other support, to raise anti-DNA antibodies using
35 DNA immunization techniques, and as an antigen to elicit an immune response.

Uses of the Polypeptides

Each of the polypeptides identified herein can be used in numerous ways. The following description should be considered exemplary and utilizes known techniques.

5 A polypeptide of the present invention can be used to assay protein levels in a biological sample using antibody-based techniques. For example, protein expression in tissues can be studied with classical immunohistological methods. (Jalkanen, M., et al., J. Cell. Biol. 101:976-985 (1985); Jalkanen, M., et al., J. Cell . Biol. 105:3087-3096 (1987).) Other antibody-based methods useful for detecting protein gene expression include immunoassays, such as the enzyme linked immunosorbent assay
10 (ELISA) and the radioimmunoassay (RIA). Suitable antibody assay labels are known in the art and include enzyme labels, such as, glucose oxidase, and radioisotopes, such as iodine (125I, 121I), carbon (14C), sulfur (35S), tritium (3H), indium (112In), and technetium (99mTc), and fluorescent labels, such as fluorescein and rhodamine, and biotin.

15 In addition to assaying secreted protein levels in a biological sample, proteins can also be detected in vivo by imaging. Antibody labels or markers for in vivo imaging of protein include those detectable by X-radiography, NMR or ESR. For X-radiography, suitable labels include radioisotopes such as barium or cesium, which emit detectable radiation but are not overtly harmful to the subject. Suitable markers for
20 NMR and ESR include those with a detectable characteristic spin, such as deuterium, which may be incorporated into the antibody by labeling of nutrients for the relevant hybridoma.

A protein-specific antibody or antibody fragment which has been labeled with an appropriate detectable imaging moiety, such as a radioisotope (for example, 131I, 25 112In, 99mTc), a radio-opaque substance, or a material detectable by nuclear magnetic resonance, is introduced (for example, parenterally, subcutaneously, or intraperitoneally) into the mammal. It will be understood in the art that the size of the subject and the imaging system used will determine the quantity of imaging moiety needed to produce diagnostic images. In the case of a radioisotope moiety, for a human
30 subject, the quantity of radioactivity injected will normally range from about 5 to 20 millicuries of 99mTc. The labeled antibody or antibody fragment will then preferentially accumulate at the location of cells which contain the specific protein. In vivo tumor imaging is described in S.W. Burchiel et al., "Immunopharmacokinetics of Radiolabeled Antibodies and Their Fragments." (Chapter 13 in Tumor Imaging: The Radiochemical Detection of Cancer, S.W. Burchiel and B. A. Rhodes, eds., Masson
35 Publishing Inc. (1982).)

Thus, the invention provides a diagnostic method of a disorder, which involves (a) assaying the expression of a polypeptide of the present invention in cells or body fluid of an individual; (b) comparing the level of gene expression with a standard gene expression level, whereby an increase or decrease in the assayed polypeptide gene
5 expression level compared to the standard expression level is indicative of a disorder.

Moreover, polypeptides of the present invention can be used to treat disease. For example, patients can be administered a polypeptide of the present invention in an effort to replace absent or decreased levels of the polypeptide (e.g., insulin), to supplement absent or decreased levels of a different polypeptide (e.g., hemoglobin S
10 for hemoglobin B), to inhibit the activity of a polypeptide (e.g., an oncogene), to activate the activity of a polypeptide (e.g., by binding to a receptor), to reduce the activity of a membrane bound receptor by competing with it for free ligand (e.g., soluble TNF receptors used in reducing inflammation), or to bring about a desired response (e.g., blood vessel growth).

15 Similarly, antibodies directed to a polypeptide of the present invention can also be used to treat disease. For example, administration of an antibody directed to a polypeptide of the present invention can bind and reduce overproduction of the polypeptide. Similarly, administration of an antibody can activate the polypeptide, such as by binding to a polypeptide bound to a membrane (receptor).

20 At the very least, the polypeptides of the present invention can be used as molecular weight markers on SDS-PAGE gels or on molecular sieve gel filtration columns using methods well known to those of skill in the art. Polypeptides can also be used to raise antibodies, which in turn are used to measure protein expression from a recombinant cell, as a way of assessing transformation of the host cell. Moreover, the
25 polypeptides of the present invention can be used to test the following biological activities.

Biological Activities

30 The polynucleotides and polypeptides of the present invention can be used in assays to test for one or more biological activities. If these polynucleotides and polypeptides do exhibit activity in a particular assay, it is likely that these molecules may be involved in the diseases associated with the biological activity. Thus, the polynucleotides and polypeptides could be used to treat the associated disease.

Immune Activity

35 A polypeptide or polynucleotide of the present invention may be useful in treating deficiencies or disorders of the immune system, by activating or inhibiting the

proliferation, differentiation, or mobilization (chemotaxis) of immune cells. Immune cells develop through a process called hematopoiesis, producing myeloid (platelets, red blood cells, neutrophils, and macrophages) and lymphoid (B and T lymphocytes) cells from pluripotent stem cells. The etiology of these immune deficiencies or disorders
5 may be genetic, somatic, such as cancer or some autoimmune disorders, acquired (e.g., by chemotherapy or toxins), or infectious. Moreover, a polynucleotide or polypeptide of the present invention can be used as a marker or detector of a particular immune system disease or disorder.

A polynucleotide or polypeptide of the present invention may be useful in
10 treating or detecting deficiencies or disorders of hematopoietic cells. A polypeptide or polynucleotide of the present invention could be used to increase differentiation and proliferation of hematopoietic cells, including the pluripotent stem cells, in an effort to treat those disorders associated with a decrease in certain (or many) types hematopoietic cells. Examples of immunologic deficiency syndromes include, but are not limited to:
15 blood protein disorders (e.g. agammaglobulinemia, dysgammaglobulinemia), ataxia telangiectasia, common variable immunodeficiency, Digeorge Syndrome, HIV infection, HTLV-BLV infection, leukocyte adhesion deficiency syndrome, lymphopenia, phagocyte bactericidal dysfunction, severe combined immunodeficiency (SCIDs), Wiskott-Aldrich Disorder, anemia, thrombocytopenia, or hemoglobinuria.

Moreover, a polypeptide or polynucleotide of the present invention could also
20 be used to modulate hemostatic (the stopping of bleeding) or thrombolytic activity (clot formation). For example, by increasing hemostatic or thrombolytic activity, a polynucleotide or polypeptide of the present invention could be used to treat blood coagulation disorders (e.g., afibrinogenemia, factor deficiencies), blood platelet
25 disorders (e.g. thrombocytopenia), or wounds resulting from trauma, surgery, or other causes. Alternatively, a polynucleotide or polypeptide of the present invention that can decrease hemostatic or thrombolytic activity could be used to inhibit or dissolve clotting. These molecules could be important in the treatment of heart attacks (infarction), strokes, or scarring.

A polynucleotide or polypeptide of the present invention may also be useful in
30 treating or detecting autoimmune disorders. Many autoimmune disorders result from inappropriate recognition of self as foreign material by immune cells. This inappropriate recognition results in an immune response leading to the destruction of the host tissue. Therefore, the administration of a polypeptide or polynucleotide of the
35 present invention that inhibits an immune response, particularly the proliferation, differentiation, or chemotaxis of T-cells, may be an effective therapy in preventing autoimmune disorders.

Examples of autoimmune disorders that can be treated or detected by the present invention include, but are not limited to: Addison's Disease, hemolytic anemia, antiphospholipid syndrome, rheumatoid arthritis, dermatitis, allergic encephalomyelitis, glomerulonephritis, Goodpasture's Syndrome, Graves' Disease, Multiple Sclerosis, 5 Myasthenia Gravis, Neuritis, Ophthalmia, Bullous Pemphigoid, Pemphigus, Polyendocrinopathies, Purpura, Reiter's Disease, Stiff-Man Syndrome, Autoimmune Thyroiditis, Systemic Lupus Erythematosus, Autoimmune Pulmonary Inflammation, Guillain-Barre Syndrome, insulin dependent diabetes mellitus, and autoimmune inflammatory eye disease.

10 Similarly, allergic reactions and conditions, such as asthma (particularly allergic asthma) or other respiratory problems, may also be treated by a polypeptide or polynucleotide of the present invention. Moreover, these molecules can be used to treat anaphylaxis, hypersensitivity to an antigenic molecule, or blood group incompatibility.

A polynucleotide or polypeptide of the present invention may also be used to 15 treat and/or prevent organ rejection or graft-versus-host disease (GVHD). Organ rejection occurs by host immune cell destruction of the transplanted tissue through an immune response. Similarly, an immune response is also involved in GVHD, but, in this case, the foreign transplanted immune cells destroy the host tissues. The administration of a polypeptide or polynucleotide of the present invention that inhibits 20 an immune response, particularly the proliferation, differentiation, or chemotaxis of T-cells, may be an effective therapy in preventing organ rejection or GVHD.

Similarly, a polypeptide or polynucleotide of the present invention may also be used to modulate inflammation. For example, the polypeptide or polynucleotide may inhibit the proliferation and differentiation of cells involved in an inflammatory 25 response. These molecules can be used to treat inflammatory conditions, both chronic and acute conditions, including inflammation associated with infection (e.g., septic shock, sepsis, or systemic inflammatory response syndrome (SIRS)), ischemia-reperfusion injury, endotoxin lethality, arthritis, complement-mediated hyperacute rejection, nephritis, cytokine or chemokine induced lung injury, inflammatory bowel 30 disease, Crohn's disease, or resulting from over production of cytokines (e.g., TNF or IL-1.)

Hyperproliferative Disorders

A polypeptide or polynucleotide can be used to treat or detect hyperproliferative 35 disorders, including neoplasms. A polypeptide or polynucleotide of the present invention may inhibit the proliferation of the disorder through direct or indirect

interactions. Alternatively, a polypeptide or polynucleotide of the present invention may proliferate other cells which can inhibit the hyperproliferative disorder.

For example, by increasing an immune response, particularly increasing antigenic qualities of the hyperproliferative disorder or by proliferating, differentiating, or mobilizing T-cells, hyperproliferative disorders can be treated. This immune response may be increased by either enhancing an existing immune response, or by initiating a new immune response. Alternatively, decreasing an immune response may also be a method of treating hyperproliferative disorders, such as a chemotherapeutic agent.

Examples of hyperproliferative disorders that can be treated or detected by a polynucleotide or polypeptide of the present invention include, but are not limited to neoplasms located in the: abdomen, bone, breast, digestive system, liver, pancreas, peritoneum, endocrine glands (adrenal, parathyroid, pituitary, testicles, ovary, thymus, thyroid), eye, head and neck, nervous (central and peripheral), lymphatic system, pelvic, skin, soft tissue, spleen, thoracic, and urogenital.

Similarly, other hyperproliferative disorders can also be treated or detected by a polynucleotide or polypeptide of the present invention. Examples of such hyperproliferative disorders include, but are not limited to: hypergammaglobulinemia, lymphoproliferative disorders, paraproteinemias, purpura, sarcoidosis, Sezary Syndrome, Waldenstrom's Macroglobulinemia, Gaucher's Disease, histiocytosis, and any other hyperproliferative disease, besides neoplasia, located in an organ system listed above.

Infectious Disease

A polypeptide or polynucleotide of the present invention can be used to treat or detect infectious agents. For example, by increasing the immune response, particularly increasing the proliferation and differentiation of B and/or T cells, infectious diseases may be treated. The immune response may be increased by either enhancing an existing immune response, or by initiating a new immune response. Alternatively, the polypeptide or polynucleotide of the present invention may also directly inhibit the infectious agent, without necessarily eliciting an immune response.

Viruses are one example of an infectious agent that can cause disease or symptoms that can be treated or detected by a polynucleotide or polypeptide of the present invention. Examples of viruses, include, but are not limited to the following DNA and RNA viral families: Arbovirus, Adenoviridae, Arenaviridae, Arterivirus, Birnaviridae, Bunyaviridae, Caliciviridae, Circoviridae, Coronaviridae, Flaviviridae, Hepadnaviridae (Hepatitis), Herpesviridae (such as, Cytomegalovirus, Herpes

Simplex, Herpes Zoster), Mononegavirus (e.g., Paramyxoviridae, Morbillivirus, Rhabdoviridae), Orthomyxoviridae (e.g., Influenza), Papovaviridae, Parvoviridae, Picornaviridae, Poxviridae (such as Smallpox or Vaccinia), Reoviridae (e.g., Rotavirus), Retroviridae (HTLV-I, HTLV-II, Lentivirus), and Togaviridae (e.g., Rubivirus). Viruses falling within these families can cause a variety of diseases or symptoms, including, but not limited to: arthritis, bronchiolitis, encephalitis, eye infections (e.g., conjunctivitis, keratitis), chronic fatigue syndrome, hepatitis (A, B, C, E, Chronic Active, Delta), meningitis, opportunistic infections (e.g., AIDS), pneumonia, Burkitt's Lymphoma, chickenpox, hemorrhagic fever, Measles, Mumps, Parainfluenza, Rabies, the common cold, Polio, leukemia, Rubella, sexually transmitted diseases, skin diseases (e.g., Kaposi's, warts), and viremia. A polypeptide or polynucleotide of the present invention can be used to treat or detect any of these symptoms or diseases.

Similarly, bacterial or fungal agents that can cause disease or symptoms and that can be treated or detected by a polynucleotide or polypeptide of the present invention include, but not limited to, the following Gram-Negative and Gram-positive bacterial families and fungi: Actinomycetales (e.g., Corynebacterium, Mycobacterium, Norcardia), Aspergillosis, Bacillaceae (e.g., Anthrax, Clostridium), Bacteroidaceae, Blastomycosis, Bordetella, Borrelia, Brucellosis, Candidiasis, Campylobacter, Coccidioidomycosis, Cryptococcosis, Dermatocycoses, Enterobacteriaceae (Klebsiella, Salmonella, Serratia, Yersinia), Erysipelothrix, Helicobacter, Legionellosis, Leptospirosis, Listeria, Mycoplasmatales, Neisseriaceae (e.g., Acinetobacter, Gonorrhea, Meningococcal), Pasteurellaceae Infections (e.g., Actinobacillus, Haemophilus, Pasteurella), Pseudomonas, Rickettsiaceae, Chlamydiaceae, Syphilis, and Staphylococcal. These bacterial or fungal families can cause the following diseases or symptoms, including, but not limited to: bacteremia, endocarditis, eye infections (conjunctivitis, tuberculosis, uveitis), gingivitis, opportunistic infections (e.g., AIDS related infections), paronychia, prosthesis-related infections, Reiter's Disease, respiratory tract infections, such as Whooping Cough or Empyema, sepsis, Lyme Disease, Cat-Scratch Disease, Dysentery, Paratyphoid Fever, food poisoning, Typhoid, pneumonia, Gonorrhea, meningitis, Chlamydia, Syphilis, Diphtheria, Leprosy, Paratuberculosis, Tuberculosis, Lupus, Botulism, gangrene, tetanus, impetigo, Rheumatic Fever, Scarlet Fever, sexually transmitted diseases, skin diseases (e.g., cellulitis, dermatocycoses), toxemia, urinary tract infections, wound infections. A polypeptide or polynucleotide of the present invention can be used to treat or detect any of these symptoms or diseases.

Moreover, parasitic agents causing disease or symptoms that can be treated or detected by a polynucleotide or polypeptide of the present invention include, but not limited to, the following families: Amebiasis, Babesiosis, Coccidiosis, Cryptosporidiosis, Dientamoebiasis, Dourine, Ectoparasitic, Giardiasis, Helminthiasis, Leishmaniasis, Theileriasis, Toxoplasmosis, Trypanosomiasis, and Trichomonas. These parasites can cause a variety of diseases or symptoms, including, but not limited to: Scabies, Trombiculiasis, eye infections, intestinal disease (e.g., dysentery, giardiasis), liver disease, lung disease, opportunistic infections (e.g., AIDS related), Malaria, pregnancy complications, and toxoplasmosis. A polypeptide or polynucleotide of the present invention can be used to treat or detect any of these symptoms or diseases.

Preferably, treatment using a polypeptide or polynucleotide of the present invention could either be by administering an effective amount of a polypeptide to the patient, or by removing cells from the patient, supplying the cells with a polynucleotide of the present invention, and returning the engineered cells to the patient (ex vivo therapy). Moreover, the polypeptide or polynucleotide of the present invention can be used as an antigen in a vaccine to raise an immune response against infectious disease.

Regeneration

A polynucleotide or polypeptide of the present invention can be used to differentiate, proliferate, and attract cells, leading to the regeneration of tissues. (See, Science 276:59-87 (1997).) The regeneration of tissues could be used to repair, replace, or protect tissue damaged by congenital defects, trauma (wounds, burns, incisions, or ulcers), age, disease (e.g. osteoporosis, osteoarthritis, periodontal disease, liver failure), surgery, including cosmetic plastic surgery, fibrosis, reperfusion injury, or systemic cytokine damage.

Tissues that could be regenerated using the present invention include organs (e.g., pancreas, liver, intestine, kidney, skin, endothelium), muscle (smooth, skeletal or cardiac), vascular (including vascular endothelium), nervous, hematopoietic, and skeletal (bone, cartilage, tendon, and ligament) tissue. Preferably, regeneration occurs without or decreased scarring. Regeneration also may include angiogenesis.

Moreover, a polynucleotide or polypeptide of the present invention may increase regeneration of tissues difficult to heal. For example, increased tendon/ligament regeneration would quicken recovery time after damage. A polynucleotide or polypeptide of the present invention could also be used prophylactically in an effort to avoid damage. Specific diseases that could be treated include of tendinitis, carpal tunnel syndrome, and other tendon or ligament defects. A further example of tissue

regeneration of non-healing wounds includes pressure ulcers, ulcers associated with vascular insufficiency, surgical, and traumatic wounds.

Similarly, nerve and brain tissue could also be regenerated by using a polynucleotide or polypeptide of the present invention to proliferate and differentiate nerve cells. Diseases that could be treated using this method include central and peripheral nervous system diseases, neuropathies, or mechanical and traumatic disorders (e.g., spinal cord disorders, head trauma, cerebrovascular disease, and stroke). Specifically, diseases associated with peripheral nerve injuries, peripheral neuropathy (e.g., resulting from chemotherapy or other medical therapies), localized neuropathies, and central nervous system diseases (e.g., Alzheimer's disease, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager syndrome), could all be treated using the polynucleotide or polypeptide of the present invention.

15 **Chemotaxis**

A polynucleotide or polypeptide of the present invention may have chemotaxis activity. A chemotactic molecule attracts or mobilizes cells (e.g., monocytes, fibroblasts, neutrophils, T-cells, mast cells, eosinophils, epithelial and/or endothelial cells) to a particular site in the body, such as inflammation, infection, or site of hyperproliferation. The mobilized cells can then fight off and/or heal the particular trauma or abnormality.

A polynucleotide or polypeptide of the present invention may increase chemotactic activity of particular cells. These chemotactic molecules can then be used to treat inflammation, infection, hyperproliferative disorders, or any immune system disorder by increasing the number of cells targeted to a particular location in the body. For example, chemotactic molecules can be used to treat wounds and other trauma to tissues by attracting immune cells to the injured location. Chemotactic molecules of the present invention can also attract fibroblasts, which can be used to treat wounds.

It is also contemplated that a polynucleotide or polypeptide of the present invention may inhibit chemotactic activity. These molecules could also be used to treat disorders. Thus, a polynucleotide or polypeptide of the present invention could be used as an inhibitor of chemotaxis.

Binding Activity

35 A polypeptide of the present invention may be used to screen for molecules that bind to the polypeptide or for molecules to which the polypeptide binds. The binding of the polypeptide and the molecule may activate (agonist), increase, inhibit

(antagonist), or decrease activity of the polypeptide or the molecule bound. Examples of such molecules include antibodies, oligonucleotides, proteins (e.g., receptors), or small molecules.

5 Preferably, the molecule is closely related to the natural ligand of the polypeptide, e.g., a fragment of the ligand, or a natural substrate, a ligand, a structural or functional mimetic. (See, Coligan et al., Current Protocols in Immunology 1(2):Chapter 5 (1991).) Similarly, the molecule can be closely related to the natural receptor to which the polypeptide binds, or at least, a fragment of the receptor capable of being bound by the polypeptide (e.g., active site). In either case, the molecule can be rationally designed using known techniques.

10 Preferably, the screening for these molecules involves producing appropriate cells which express the polypeptide, either as a secreted protein or on the cell membrane. Preferred cells include cells from mammals, yeast, *Drosophila*, or *E. coli*. Cells expressing the polypeptide (or cell membrane containing the expressed polypeptide) are then preferably contacted with a test compound potentially containing the molecule to observe binding, stimulation, or inhibition of activity of either the polypeptide or the molecule.

The assay may simply test binding of a candidate compound to the polypeptide, wherein binding is detected by a label, or in an assay involving competition with a labeled competitor. Further, the assay may test whether the candidate compound results in a signal generated by binding to the polypeptide.

20 Alternatively, the assay can be carried out using cell-free preparations, polypeptide/molecule affixed to a solid support, chemical libraries, or natural product mixtures. The assay may also simply comprise the steps of mixing a candidate compound with a solution containing a polypeptide, measuring polypeptide/molecule activity or binding, and comparing the polypeptide/molecule activity or binding to a standard.

25 Preferably, an ELISA assay can measure polypeptide level or activity in a sample (e.g., biological sample) using a monoclonal or polyclonal antibody. The antibody can measure polypeptide level or activity by either binding, directly or indirectly, to the polypeptide or by competing with the polypeptide for a substrate.

30 All of these above assays can be used as diagnostic or prognostic markers. The molecules discovered using these assays can be used to treat disease or to bring about a particular result in a patient (e.g., blood vessel growth) by activating or inhibiting the polypeptide/molecule. Moreover, the assays can discover agents which may inhibit or enhance the production of the polypeptide from suitably manipulated cells or tissues.

Therefore, the invention includes a method of identifying compounds which bind to a polypeptide of the invention comprising the steps of: (a) incubating a candidate binding compound with a polypeptide of the invention; and (b) determining if binding has occurred. Moreover, the invention includes a method of identifying agonists/antagonists comprising the steps of: (a) incubating a candidate compound with a polypeptide of the invention, (b) assaying a biological activity, and (b) determining if a biological activity of the polypeptide has been altered.

Other Activities

10 A polypeptide or polynucleotide of the present invention may also increase or decrease the differentiation or proliferation of embryonic stem cells, besides, as discussed above, hematopoietic lineage.

A polypeptide or polynucleotide of the present invention may also be used to modulate mammalian characteristics, such as body height, weight, hair color, eye color, skin, percentage of adipose tissue, pigmentation, size, and shape (e.g., cosmetic surgery). Similarly, a polypeptide or polynucleotide of the present invention may be used to modulate mammalian metabolism affecting catabolism, anabolism, processing, utilization, and storage of energy.

20 A polypeptide or polynucleotide of the present invention may be used to change a mammal's mental state or physical state by influencing biorhythms, circadian rhythms, depression (including depressive disorders), tendency for violence, tolerance for pain, reproductive capabilities (preferably by Activin or Inhibin-like activity), hormonal or endocrine levels, appetite, libido, memory, stress, or other cognitive qualities.

25 A polypeptide or polynucleotide of the present invention may also be used as a food additive or preservative, such as to increase or decrease storage capabilities, fat content, lipid, protein, carbohydrate, vitamins, minerals, cofactors or other nutritional components.

30 Other Preferred Embodiments

Other preferred embodiments of the claimed invention include an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a sequence of at least about 50 contiguous nucleotides in the nucleotide sequence of SEQ ID NO:X wherein X is any integer as defined in Table 1.

35 Also preferred is a nucleic acid molecule wherein said sequence of contiguous nucleotides is included in the nucleotide sequence of SEQ ID NO:X in the range of

positions beginning with the nucleotide at about the position of the 5' Nucleotide of the Clone Sequence and ending with the nucleotide at about the position of the 3' Nucleotide of the Clone Sequence as defined for SEQ ID NO:X in Table 1.

Also preferred is a nucleic acid molecule wherein said sequence of contiguous
5 nucleotides is included in the nucleotide sequence of SEQ ID NO:X in the range of positions beginning with the nucleotide at about the position of the 5' Nucleotide of the Start Codon and ending with the nucleotide at about the position of the 3' Nucleotide of the Clone Sequence as defined for SEQ ID NO:X in Table 1.

Similarly preferred is a nucleic acid molecule wherein said sequence of
10 contiguous nucleotides is included in the nucleotide sequence of SEQ ID NO:X in the range of positions beginning with the nucleotide at about the position of the 5' Nucleotide of the First Amino Acid of the Signal Peptide and ending with the nucleotide at about the position of the 3' Nucleotide of the Clone Sequence as defined for SEQ ID NO:X in Table 1.

Also preferred is an isolated nucleic acid molecule comprising a nucleotide
15 sequence which is at least 95% identical to a sequence of at least about 150 contiguous nucleotides in the nucleotide sequence of SEQ ID NO:X.

Further preferred is an isolated nucleic acid molecule comprising a nucleotide
20 sequence which is at least 95% identical to a sequence of at least about 500 contiguous nucleotides in the nucleotide sequence of SEQ ID NO:X.

A further preferred embodiment is a nucleic acid molecule comprising a
nucleotide sequence which is at least 95% identical to the nucleotide sequence of SEQ
ID NO:X beginning with the nucleotide at about the position of the 5' Nucleotide of the
First Amino Acid of the Signal Peptide and ending with the nucleotide at about the
25 position of the 3' Nucleotide of the Clone Sequence as defined for SEQ ID NO:X in
Table 1.

A further preferred embodiment is an isolated nucleic acid molecule comprising
a nucleotide sequence which is at least 95% identical to the complete nucleotide
sequence of SEQ ID NO:X.

Also preferred is an isolated nucleic acid molecule which hybridizes under
30 stringent hybridization conditions to a nucleic acid molecule, wherein said nucleic acid molecule which hybridizes does not hybridize under stringent hybridization conditions to a nucleic acid molecule having a nucleotide sequence consisting of only A residues or of only T residues.

Also preferred is a composition of matter comprising a DNA molecule which
35 comprises a human cDNA clone identified by a cDNA Clone Identifier in Table 1, which DNA molecule is contained in the material deposited with the American Type

Culture Collection and given the ATCC Deposit Number shown in Table 1 for said cDNA Clone Identifier.

Also preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a sequence of at least 50 contiguous nucleotides in the nucleotide sequence of a human cDNA clone identified by a cDNA Clone Identifier in Table 1, which DNA molecule is contained in the deposit given the ATCC Deposit Number shown in Table 1.

Also preferred is an isolated nucleic acid molecule, wherein said sequence of at least 50 contiguous nucleotides is included in the nucleotide sequence of the complete open reading frame sequence encoded by said human cDNA clone.

Also preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to sequence of at least 150 contiguous nucleotides in the nucleotide sequence encoded by said human cDNA clone.

A further preferred embodiment is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to sequence of at least 500 contiguous nucleotides in the nucleotide sequence encoded by said human cDNA clone.

A further preferred embodiment is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to the complete nucleotide sequence encoded by said human cDNA clone.

A further preferred embodiment is a method for detecting in a biological sample a nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from the group consisting of: a nucleotide sequence of SEQ ID NO:X wherein X is any integer as defined in Table 1; and a nucleotide sequence encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1; which method comprises a step of comparing a nucleotide sequence of at least one nucleic acid molecule in said sample with a sequence selected from said group and determining whether the sequence of said nucleic acid molecule in said sample is at least 95% identical to said selected sequence.

Also preferred is the above method wherein said step of comparing sequences comprises determining the extent of nucleic acid hybridization between nucleic acid molecules in said sample and a nucleic acid molecule comprising said sequence selected from said group. Similarly, also preferred is the above method wherein said step of comparing sequences is performed by comparing the nucleotide sequence determined from a nucleic acid molecule in said sample with said sequence selected from said group. The nucleic acid molecules can comprise DNA molecules or RNA molecules.

A further preferred embodiment is a method for identifying the species, tissue or cell type of a biological sample which method comprises a step of detecting nucleic acid molecules in said sample, if any, comprising a nucleotide sequence that is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from the group consisting of: a nucleotide sequence of SEQ ID NO:X wherein X is any integer as defined in Table 1; and a nucleotide sequence encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

The method for identifying the species, tissue or cell type of a biological sample can comprise a step of detecting nucleic acid molecules comprising a nucleotide sequence in a panel of at least two nucleotide sequences, wherein at least one sequence in said panel is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from said group.

Also preferred is a method for diagnosing in a subject a pathological condition associated with abnormal structure or expression of a gene encoding a secreted protein identified in Table 1, which method comprises a step of detecting in a biological sample obtained from said subject nucleic acid molecules, if any, comprising a nucleotide sequence that is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from the group consisting of: a nucleotide sequence of SEQ ID NO:X wherein X is any integer as defined in Table 1; and a nucleotide sequence encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

The method for diagnosing a pathological condition can comprise a step of detecting nucleic acid molecules comprising a nucleotide sequence in a panel of at least two nucleotide sequences, wherein at least one sequence in said panel is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from said group.

Also preferred is a composition of matter comprising isolated nucleic acid molecules wherein the nucleotide sequences of said nucleic acid molecules comprise a panel of at least two nucleotide sequences, wherein at least one sequence in said panel is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from the group consisting of: a nucleotide sequence of SEQ ID NO:X wherein X is any integer as defined in Table 1; and a nucleotide sequence encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1. The nucleic acid molecules can comprise DNA molecules or RNA molecules.

Also preferred is an isolated polypeptide comprising an amino acid sequence at least 90% identical to a sequence of at least about 10 contiguous amino acids in the amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1.

Also preferred is a polypeptide, wherein said sequence of contiguous amino acids is included in the amino acid sequence of SEQ ID NO:Y in the range of positions beginning with the residue at about the position of the First Amino Acid of the Secreted Portion and ending with the residue at about the Last Amino Acid of the Open Reading Frame as set forth for SEQ ID NO:Y in Table 1.

Also preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence of at least about 30 contiguous amino acids in the amino acid sequence of SEQ ID NO:Y.

Further preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence of at least about 100 contiguous amino acids in the amino acid sequence of SEQ ID NO:Y.

Further preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to the complete amino acid sequence of SEQ ID NO:Y.

Further preferred is an isolated polypeptide comprising an amino acid sequence at least 90% identical to a sequence of at least about 10 contiguous amino acids in the complete amino acid sequence of a secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is a polypeptide wherein said sequence of contiguous amino acids is included in the amino acid sequence of a secreted portion of the secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence of at least about 30 contiguous amino acids in the amino acid sequence of the secreted portion of the protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence of at least about 100 contiguous amino acids in the amino acid sequence of the secreted portion of the protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to the amino acid sequence of the secreted portion of the protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Further preferred is an isolated antibody which binds specifically to a polypeptide comprising an amino acid sequence that is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Further preferred is a method for detecting in a biological sample a polypeptide comprising an amino acid sequence which is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1; which method comprises a step of comparing an amino acid sequence of at least one polypeptide molecule in said sample with a sequence selected from said group and determining whether the sequence of said polypeptide molecule in said sample is at least 90% identical to said sequence of at least 10 contiguous amino acids.

Also preferred is the above method wherein said step of comparing an amino acid sequence of at least one polypeptide molecule in said sample with a sequence selected from said group comprises determining the extent of specific binding of polypeptides in said sample to an antibody which binds specifically to a polypeptide comprising an amino acid sequence that is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is the above method wherein said step of comparing sequences is performed by comparing the amino acid sequence determined from a polypeptide molecule in said sample with said sequence selected from said group.

Also preferred is a method for identifying the species, tissue or cell type of a biological sample which method comprises a step of detecting polypeptide molecules in said sample, if any, comprising an amino acid sequence that is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is the above method for identifying the species, tissue or cell type of a biological sample, which method comprises a step of detecting polypeptide molecules comprising an amino acid sequence in a panel of at least two amino acid sequences, wherein at least one sequence in said panel is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the above group.

Also preferred is a method for diagnosing in a subject a pathological condition associated with abnormal structure or expression of a gene encoding a secreted protein identified in Table 1, which method comprises a step of detecting in a biological sample obtained from said subject polypeptide molecules comprising an amino acid sequence in a panel of at least two amino acid sequences, wherein at least one sequence in said panel is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

In any of these methods, the step of detecting said polypeptide molecules includes using an antibody.

Also preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a nucleotide sequence encoding a polypeptide wherein said polypeptide comprises an amino acid sequence that is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is an isolated nucleic acid molecule, wherein said nucleotide sequence encoding a polypeptide has been optimized for expression of said polypeptide in a prokaryotic host.

5 Also preferred is an isolated nucleic acid molecule, wherein said polypeptide comprises an amino acid sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

10 Further preferred is a method of making a recombinant vector comprising inserting any of the above isolated nucleic acid molecule into a vector. Also preferred is the recombinant vector produced by this method. Also preferred is a method of making a recombinant host cell comprising introducing the vector into a host cell, as well as the recombinant host cell produced by this method.

15 Also preferred is a method of making an isolated polypeptide comprising culturing this recombinant host cell under conditions such that said polypeptide is expressed and recovering said polypeptide. Also preferred is this method of making an isolated polypeptide, wherein said recombinant host cell is a eukaryotic cell and said polypeptide is a secreted portion of a human secreted protein comprising an amino acid
20 sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y beginning with the residue at the position of the First Amino Acid of the Secreted Portion of SEQ ID NO:Y wherein Y is an integer set forth in Table 1 and said position of the First Amino Acid of the Secreted Portion of SEQ ID NO:Y is defined in Table 1; and an amino acid sequence of a secreted portion of a protein encoded by a human
25 cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1. The isolated polypeptide produced by this method is also preferred.

Also preferred is a method of treatment of an individual in need of an increased level of a secreted protein activity, which method comprises administering to such an
30 individual a pharmaceutical composition comprising an amount of an isolated polypeptide, polynucleotide, or antibody of the claimed invention effective to increase the level of said protein activity in said individual.

Having generally described the invention, the same will be more readily understood by reference to the following examples, which are provided by way of
35 illustration and are not intended as limiting.

Examples

Example 1: Isolation of a Selected cDNA Clone From the Deposited Sample

5 Each cDNA clone in a cited ATCC deposit is contained in a plasmid vector. Table 1 identifies the vectors used to construct the cDNA library from which each clone was isolated. In many cases, the vector used to construct the library is a phage vector from which a plasmid has been excised. The table immediately below correlates the related plasmid for each phage vector used in constructing the cDNA library. For
10 example, where a particular clone is identified in Table 1 as being isolated in the vector "Lambda Zap," the corresponding deposited clone is in "pBluescript."

	<u>Vector Used to Construct Library</u>	<u>Corresponding Deposited Plasmid</u>
	Lambda Zap	pBluescript (pBS)
	Uni-Zap XR	pBluescript (pBS)
15	Zap Express	pBK
	lafmid BA	plafmid BA
	pSport1	pSport1
	pCMVSPORT 2.0	pCMVSPORT 2.0
	pCMVSPORT 3.0	pCMVSPORT 3.0
20	pCR [®] 2.1	pCR [®] 2.1

Vectors Lambda Zap (U.S. Patent Nos. 5,128,256 and 5,286,636), Uni-Zap XR (U.S. Patent Nos. 5,128, 256 and 5,286,636), Zap Express (U.S. Patent Nos. 5,128,256 and 5,286,636), pBluescript (pBS) (Short, J. M. et al., Nucleic Acids Res. 16:7583-7600 (1988); Alting-Mees, M. A. and Short, J. M., Nucleic Acids Res. 25 17:9494 (1989)) and pBK (Alting-Mees, M. A. et al., Strategies 5:58-61 (1992)) are commercially available from Stratagene Cloning Systems, Inc., 11011 N. Torrey Pines Road, La Jolla, CA, 92037. pBS contains an ampicillin resistance gene and pBK contains a neomycin resistance gene. Both can be transformed into E. coli strain XL-1 Blue, also available from Stratagene. pBS comes in 4 forms SK+, SK-, KS+ and KS. The S and K refers to the orientation of the polylinker to the T7 and T3 primer
30 sequences which flank the polylinker region ("S" is for SacI and "K" is for KpnI which are the first sites on each respective end of the linker). "+" or "-" refer to the orientation of the f1 origin of replication ("ori"), such that in one orientation, single stranded rescue initiated from the f1 ori generates sense strand DNA and in the other, antisense.

35 Vectors pSport1, pCMVSPORT 2.0 and pCMVSPORT 3.0, were obtained from Life Technologies, Inc., P. O. Box 6009, Gaithersburg, MD 20897. All Sport vectors contain an ampicillin resistance gene and may be transformed into E. coli strain

DH10B, also available from Life Technologies. (See, for instance, Gruber, C. E., et al., *Focus* 15:59 (1993).) Vector lafmid BA (Bento Soares, Columbia University, NY) contains an ampicillin resistance gene and can be transformed into *E. coli* strain XL-1 Blue. Vector pCR[®]2.1, which is available from Invitrogen, 1600 Faraday Avenue, Carlsbad, CA 92008, contains an ampicillin resistance gene and may be transformed into *E. coli* strain DH10B, available from Life Technologies. (See, for instance, Clark, J. M., *Nuc. Acids Res.* 16:9677-9686 (1988) and Mead, D. et al., *Bio/Technology* 9: (1991).) Preferably, a polynucleotide of the present invention does not comprise the phage vector sequences identified for the particular clone in Table 1, as well as the corresponding plasmid vector sequences designated above.

The deposited material in the sample assigned the ATCC Deposit Number cited in Table 1 for any given cDNA clone also may contain one or more additional plasmids, each comprising a cDNA clone different from that given clone. Thus, deposits sharing the same ATCC Deposit Number contain at least a plasmid for each cDNA clone identified in Table 1. Typically, each ATCC deposit sample cited in Table 1 comprises a mixture of approximately equal amounts (by weight) of about 50 plasmid DNAs, each containing a different cDNA clone; but such a deposit sample may include plasmids for more or less than 50 cDNA clones, up to about 500 cDNA clones.

Two approaches can be used to isolate a particular clone from the deposited sample of plasmid DNAs cited for that clone in Table 1. First, a plasmid is directly isolated by screening the clones using a polynucleotide probe corresponding to SEQ ID NO:X.

Particularly, a specific polynucleotide with 30-40 nucleotides is synthesized using an Applied Biosystems DNA synthesizer according to the sequence reported. The oligonucleotide is labeled, for instance, with ³²P-γ-ATP using T4 polynucleotide kinase and purified according to routine methods. (E.g., Maniatis et al., *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Press, Cold Spring, NY (1982).) The plasmid mixture is transformed into a suitable host, as indicated above (such as XL-1 Blue (Stratagene)) using techniques known to those of skill in the art, such as those provided by the vector supplier or in related publications or patents cited above. The transformants are plated on 1.5% agar plates (containing the appropriate selection agent, e.g., ampicillin) to a density of about 150 transformants (colonies) per plate. These plates are screened using Nylon membranes according to routine methods for bacterial colony screening (e.g., Sambrook et al., *Molecular Cloning: A Laboratory Manual*, 2nd Edit., (1989), Cold Spring Harbor Laboratory Press, pages 1.93 to 1.104), or other techniques known to those of skill in the art.

Alternatively, two primers of 17-20 nucleotides derived from both ends of the SEQ ID NO:X (i.e., within the region of SEQ ID NO:X bounded by the 5' NT and the 3' NT of the clone defined in Table 1) are synthesized and used to amplify the desired cDNA using the deposited cDNA plasmid as a template. The polymerase chain reaction is carried out under routine conditions, for instance, in 25 μ l of reaction mixture with 0.5 μ g of the above cDNA template. A convenient reaction mixture is 1.5-5 mM $MgCl_2$, 0.01% (w/v) gelatin, 20 μ M each of dATP, dCTP, dGTP, dTTP, 25 pmol of each primer and 0.25 Unit of Taq polymerase. Thirty five cycles of PCR (denaturation at 94°C for 1 min; annealing at 55°C for 1 min; elongation at 72°C for 1 min) are performed with a Perkin-Elmer Cetus automated thermal cycler. The amplified product is analyzed by agarose gel electrophoresis and the DNA band with expected molecular weight is excised and purified. The PCR product is verified to be the selected sequence by subcloning and sequencing the DNA product.

Several methods are available for the identification of the 5' or 3' non-coding portions of a gene which may not be present in the deposited clone. These methods include but are not limited to, filter probing, clone enrichment using specific probes, and protocols similar or identical to 5' and 3' "RACE" protocols which are well known in the art. For instance, a method similar to 5' RACE is available for generating the missing 5' end of a desired full-length transcript. (Fromont-Racine et al., *Nucleic Acids Res.* 21(7):1683-1684 (1993).)

Briefly, a specific RNA oligonucleotide is ligated to the 5' ends of a population of RNA presumably containing full-length gene RNA transcripts. A primer set containing a primer specific to the ligated RNA oligonucleotide and a primer specific to a known sequence of the gene of interest is used to PCR amplify the 5' portion of the desired full-length gene. This amplified product may then be sequenced and used to generate the full length gene.

This above method starts with total RNA isolated from the desired source, although poly-A+ RNA can be used. The RNA preparation can then be treated with phosphatase if necessary to eliminate 5' phosphate groups on degraded or damaged RNA which may interfere with the later RNA ligase step. The phosphatase should then be inactivated and the RNA treated with tobacco acid pyrophosphatase in order to remove the cap structure present at the 5' ends of messenger RNAs. This reaction leaves a 5' phosphate group at the 5' end of the cap cleaved RNA which can then be ligated to an RNA oligonucleotide using T4 RNA ligase.

This modified RNA preparation is used as a template for first strand cDNA synthesis using a gene specific oligonucleotide. The first strand synthesis reaction is

used as a template for PCR amplification of the desired 5' end using a primer specific to the ligated RNA oligonucleotide and a primer specific to the known sequence of the gene of interest. The resultant product is then sequenced and analyzed to confirm that the 5' end sequence belongs to the desired gene.

5

Example 2: Isolation of Genomic Clones Corresponding to a Polynucleotide

A human genomic P1 library (Genomic Systems, Inc.) is screened by PCR using primers selected for the cDNA sequence corresponding to SEQ ID NO:X., according to the method described in Example 1. (See also, Sambrook.)

10

Example 3: Tissue Distribution of Polypeptide

Tissue distribution of mRNA expression of polynucleotides of the present invention is determined using protocols for Northern blot analysis, described by, among others, Sambrook et al. For example, a cDNA probe produced by the method described in Example 1 is labeled with P³² using the rediprime™ DNA labeling system (Amersham Life Science), according to manufacturer's instructions. After labeling, the probe is purified using CHROMA SPIN-100™ column (Clontech Laboratories, Inc.), according to manufacturer's protocol number PT1200-1. The purified labeled probe is then used to examine various human tissues for mRNA expression.

15

20

Multiple Tissue Northern (MTN) blots containing various human tissues (H) or human immune system tissues (IM) (Clontech) are examined with the labeled probe using ExpressHyb™ hybridization solution (Clontech) according to manufacturer's protocol number PT1190-1. Following hybridization and washing, the blots are mounted and exposed to film at -70°C overnight, and the films developed according to standard procedures.

25

Example 4: Chromosomal Mapping of the Polynucleotides

An oligonucleotide primer set is designed according to the sequence at the 5' end of SEQ ID NO:X. This primer preferably spans about 100 nucleotides. This primer set is then used in a polymerase chain reaction under the following set of conditions : 30 seconds, 95°C; 1 minute, 56°C; 1 minute, 70°C. This cycle is repeated 32 times followed by one 5 minute cycle at 70°C. Human, mouse, and hamster DNA is used as template in addition to a somatic cell hybrid panel containing individual chromosomes or chromosome fragments (Bios, Inc). The reactions is analyzed on

30

35

either 8% polyacrylamide gels or 3.5 % agarose gels. Chromosome mapping is determined by the presence of an approximately 100 bp PCR fragment in the particular somatic cell hybrid.

5 **Example 5: Bacterial Expression of a Polypeptide**

 A polynucleotide encoding a polypeptide of the present invention is amplified using PCR oligonucleotide primers corresponding to the 5' and 3' ends of the DNA sequence, as outlined in Example 1, to synthesize insertion fragments. The primers used to amplify the cDNA insert should preferably contain restriction sites, such as
10 BamHI and XbaI, at the 5' end of the primers in order to clone the amplified product into the expression vector. For example, BamHI and XbaI correspond to the restriction enzyme sites on the bacterial expression vector pQE-9. (Qiagen, Inc., Chatsworth, CA). This plasmid vector encodes antibiotic resistance (Amp^r), a bacterial origin of replication (ori), an IPTG-regulatable promoter/operator (P/O), a ribosome binding site
15 (RBS), a 6-histidine tag (6-His), and restriction enzyme cloning sites.

 The pQE-9 vector is digested with BamHI and XbaI and the amplified fragment is ligated into the pQE-9 vector maintaining the reading frame initiated at the bacterial RBS. The ligation mixture is then used to transform the E. coli strain M15/rep4 (Qiagen, Inc.) which contains multiple copies of the plasmid pREP4, which expresses
20 the lacI repressor and also confers kanamycin resistance (Kan^r). Transformants are identified by their ability to grow on LB plates and ampicillin/kanamycin resistant colonies are selected. Plasmid DNA is isolated and confirmed by restriction analysis.

 Clones containing the desired constructs are grown overnight (O/N) in liquid culture in LB media supplemented with both Amp (100 ug/ml) and Kan (25 ug/ml).
25 The O/N culture is used to inoculate a large culture at a ratio of 1:100 to 1:250. The cells are grown to an optical density 600 (O.D.⁶⁰⁰) of between 0.4 and 0.6. IPTG (Isopropyl-B-D-thiogalacto pyranoside) is then added to a final concentration of 1 mM. IPTG induces by inactivating the lacI repressor, clearing the P/O leading to increased gene expression.

30 Cells are grown for an extra 3 to 4 hours. Cells are then harvested by centrifugation (20 mins at 6000Xg). The cell pellet is solubilized in the chaotropic agent 6 Molar Guanidine HCl by stirring for 3-4 hours at 4°C. The cell debris is removed by centrifugation, and the supernatant containing the polypeptide is loaded onto a nickel-nitrilo-tri-acetic acid ("Ni-NTA") affinity resin column (available from
35 QIAGEN, Inc., *supra*). Proteins with a 6 x His tag bind to the Ni-NTA resin with high

affinity and can be purified in a simple one-step procedure (for details see: The QIAexpressionist (1995) QIAGEN, Inc., *supra*).

Briefly, the supernatant is loaded onto the column in 6 M guanidine-HCl, pH 8, the column is first washed with 10 volumes of 6 M guanidine-HCl, pH 8, then washed
5 with 10 volumes of 6 M guanidine-HCl pH 6, and finally the polypeptide is eluted with 6 M guanidine-HCl, pH 5.

The purified protein is then renatured by dialyzing it against phosphate-buffered saline (PBS) or 50 mM Na-acetate, pH 6 buffer plus 200 mM NaCl. Alternatively, the protein can be successfully refolded while immobilized on the Ni-NTA column. The
10 recommended conditions are as follows: renature using a linear 6M-1M urea gradient in 500 mM NaCl, 20% glycerol, 20 mM Tris/HCl pH 7.4, containing protease inhibitors. The renaturation should be performed over a period of 1.5 hours or more. After renaturation the proteins are eluted by the addition of 250 mM imidazole. Imidazole is removed by a final dialyzing step against PBS or 50 mM sodium acetate pH 6 buffer
15 plus 200 mM NaCl. The purified protein is stored at 4°C or frozen at -80°C.

In addition to the above expression vector, the present invention further includes an expression vector comprising phage operator and promoter elements operatively linked to a polynucleotide of the present invention, called pHE4a. (ATCC Accession Number 209645, deposited on February 25, 1998.) This vector contains: 1) a
20 neomycinphosphotransferase gene as a selection marker, 2) an E. coli origin of replication, 3) a T5 phage promoter sequence, 4) two lac operator sequences, 5) a Shine-Delgarno sequence, and 6) the lactose operon repressor gene (*lacIq*). The origin of replication (*oriC*) is derived from pUC19 (LTI, Gaithersburg, MD). The promoter sequence and operator sequences are made synthetically.

DNA can be inserted into the pHEa by restricting the vector with NdeI and XbaI, BamHI, XhoI, or Asp718, running the restricted product on a gel, and isolating the larger fragment (the stuffer fragment should be about 310 base pairs). The DNA insert is generated according to the PCR protocol described in Example 1, using PCR
25 primers having restriction sites for NdeI (5' primer) and XbaI, BamHI, XhoI, or Asp718 (3' primer). The PCR insert is gel purified and restricted with compatible enzymes. The insert and vector are ligated according to standard protocols.
30

The engineered vector could easily be substituted in the above protocol to express protein in a bacterial system.

35 **Example 6: Purification of a Polypeptide from an Inclusion Body**

The following alternative method can be used to purify a polypeptide expressed in *E. coli* when it is present in the form of inclusion bodies. Unless otherwise specified, all of the following steps are conducted at 4-10°C.

Upon completion of the production phase of the *E. coli* fermentation, the cell
5 culture is cooled to 4-10°C and the cells harvested by continuous centrifugation at 15,000 rpm (Heraeus Sepatech). On the basis of the expected yield of protein per unit weight of cell paste and the amount of purified protein required, an appropriate amount of cell paste, by weight, is suspended in a buffer solution containing 100 mM Tris, 50 mM EDTA, pH 7.4. The cells are dispersed to a homogeneous suspension using a
10 high shear mixer.

The cells are then lysed by passing the solution through a microfluidizer (Microfluidics, Corp. or APV Gaulin, Inc.) twice at 4000-6000 psi. The homogenate is then mixed with NaCl solution to a final concentration of 0.5 M NaCl, followed by centrifugation at 7000 xg for 15 min. The resultant pellet is washed again using 0.5M
15 NaCl, 100 mM Tris, 50 mM EDTA, pH 7.4.

The resulting washed inclusion bodies are solubilized with 1.5 M guanidine hydrochloride (GuHCl) for 2-4 hours. After 7000 xg centrifugation for 15 min., the pellet is discarded and the polypeptide containing supernatant is incubated at 4°C overnight to allow further GuHCl extraction.

20 Following high speed centrifugation (30,000 xg) to remove insoluble particles, the GuHCl solubilized protein is refolded by quickly mixing the GuHCl extract with 20 volumes of buffer containing 50 mM sodium, pH 4.5, 150 mM NaCl, 2 mM EDTA by vigorous stirring. The refolded diluted protein solution is kept at 4°C without mixing for 12 hours prior to further purification steps.

25 To clarify the refolded polypeptide solution, a previously prepared tangential filtration unit equipped with 0.16 µm membrane filter with appropriate surface area (e.g., Filtron), equilibrated with 40 mM sodium acetate, pH 6.0 is employed. The filtered sample is loaded onto a cation exchange resin (e.g., Poros HS-50, Perseptive Biosystems). The column is washed with 40 mM sodium acetate, pH 6.0 and eluted
30 with 250 mM, 500 mM, 1000 mM, and 1500 mM NaCl in the same buffer, in a stepwise manner. The absorbance at 280 nm of the effluent is continuously monitored. Fractions are collected and further analyzed by SDS-PAGE.

Fractions containing the polypeptide are then pooled and mixed with 4 volumes of water. The diluted sample is then loaded onto a previously prepared set of tandem

columns of strong anion (Poros HQ-50, Perseptive Biosystems) and weak anion (Poros CM-20, Perseptive Biosystems) exchange resins. The columns are equilibrated with 40 mM sodium acetate, pH 6.0. Both columns are washed with 40 mM sodium acetate, pH 6.0, 200 mM NaCl. The CM-20 column is then eluted using a 10 column
5 volume linear gradient ranging from 0.2 M NaCl, 50 mM sodium acetate, pH 6.0 to 1.0 M NaCl, 50 mM sodium acetate, pH 6.5. Fractions are collected under constant A_{280} monitoring of the effluent. Fractions containing the polypeptide (determined, for instance, by 16% SDS-PAGE) are then pooled.

The resultant polypeptide should exhibit greater than 95% purity after the above
10 refolding and purification steps. No major contaminant bands should be observed from Commaissie blue stained 16% SDS-PAGE gel when 5 μ g of purified protein is loaded. The purified protein can also be tested for endotoxin/LPS contamination, and typically the LPS content is less than 0.1 ng/ml according to LAL assays.

15 **Example 7: Cloning and Expression of a Polypeptide in a Baculovirus Expression System**

In this example, the plasmid shuttle vector pA2 is used to insert a polynucleotide into a baculovirus to express a polypeptide. This expression vector contains the strong polyhedrin promoter of the *Autographa californica* nuclear polyhedrosis virus
20 (AcMNPV) followed by convenient restriction sites such as BamHI, Xba I and Asp718. The polyadenylation site of the simian virus 40 ("SV40") is used for efficient polyadenylation. For easy selection of recombinant virus, the plasmid contains the beta-galactosidase gene from *E. coli* under control of a weak *Drosophila* promoter in the same orientation, followed by the polyadenylation signal of the polyhedrin gene. The
25 inserted genes are flanked on both sides by viral sequences for cell-mediated homologous recombination with wild-type viral DNA to generate a viable virus that express the cloned polynucleotide.

Many other baculovirus vectors can be used in place of the vector above, such as pAc373, pVL941, and pAcIM1, as one skilled in the art would readily appreciate, as
30 long as the construct provides appropriately located signals for transcription, translation, secretion and the like, including a signal peptide and an in-frame AUG as required. Such vectors are described, for instance, in Luckow et al., *Virology* 170:31-39 (1989).

Specifically, the cDNA sequence contained in the deposited clone, including the
35 AUG initiation codon and the naturally associated leader sequence identified in Table 1, is amplified using the PCR protocol described in Example 1. If the naturally occurring

signal sequence is used to produce the secreted protein, the pA2 vector does not need a second signal peptide. Alternatively, the vector can be modified (pA2 GP) to include a baculovirus leader sequence, using the standard methods described in Summers et al., "A Manual of Methods for Baculovirus Vectors and Insect Cell Culture Procedures,"

5 Texas Agricultural Experimental Station Bulletin No. 1555 (1987).

The amplified fragment is isolated from a 1% agarose gel using a commercially available kit ("GeneClean," BIO 101 Inc., La Jolla, Ca.). The fragment then is digested with appropriate restriction enzymes and again purified on a 1% agarose gel.

10 The plasmid is digested with the corresponding restriction enzymes and optionally, can be dephosphorylated using calf intestinal phosphatase, using routine procedures known in the art. The DNA is then isolated from a 1% agarose gel using a commercially available kit ("GeneClean" BIO 101 Inc., La Jolla, Ca.).

The fragment and the dephosphorylated plasmid are ligated together with T4 DNA ligase. *E. coli* HB101 or other suitable *E. coli* hosts such as XL-1 Blue
15 (Stratagene Cloning Systems, La Jolla, CA) cells are transformed with the ligation mixture and spread on culture plates. Bacteria containing the plasmid are identified by digesting DNA from individual colonies and analyzing the digestion product by gel electrophoresis. The sequence of the cloned fragment is confirmed by DNA sequencing.

20 Five μ g of a plasmid containing the polynucleotide is co-transfected with 1.0 μ g of a commercially available linearized baculovirus DNA ("BaculoGold™ baculovirus DNA", Pharmingen, San Diego, CA), using the lipofection method described by Felgner et al., Proc. Natl. Acad. Sci. USA 84:7413-7417 (1987). One μ g of BaculoGold™ virus DNA and 5 μ g of the plasmid are mixed in a sterile well of a
25 microtiter plate containing 50 μ l of serum-free Grace's medium (Life Technologies Inc., Gaithersburg, MD). Afterwards, 10 μ l Lipofectin plus 90 μ l Grace's medium are added, mixed and incubated for 15 minutes at room temperature. Then the transfection mixture is added drop-wise to Sf9 insect cells (ATCC CRL 1711) seeded in a 35 mm tissue culture plate with 1 ml Grace's medium without serum. The plate is then
30 incubated for 5 hours at 27° C. The transfection solution is then removed from the plate and 1 ml of Grace's insect medium supplemented with 10% fetal calf serum is added. Cultivation is then continued at 27° C for four days.

After four days the supernatant is collected and a plaque assay is performed, as described by Summers and Smith, *supra*. An agarose gel with "Blue Gal" (Life
35 Technologies Inc., Gaithersburg) is used to allow easy identification and isolation of gal-expressing clones, which produce blue-stained plaques. (A detailed description of a "plaque assay" of this type can also be found in the user's guide for insect cell culture

and baculovirology distributed by Life Technologies Inc., Gaithersburg, page 9-10.) After appropriate incubation, blue stained plaques are picked with the tip of a micropipettor (e.g., Eppendorf). The agar containing the recombinant viruses is then resuspended in a microcentrifuge tube containing 200 μ l of Grace's medium and the suspension containing the recombinant baculovirus is used to infect Sf9 cells seeded in 35 mm dishes. Four days later the supernatants of these culture dishes are harvested and then they are stored at 4° C.

To verify the expression of the polypeptide, Sf9 cells are grown in Grace's medium supplemented with 10% heat-inactivated FBS. The cells are infected with the recombinant baculovirus containing the polynucleotide at a multiplicity of infection ("MOI") of about 2. If radiolabeled proteins are desired, 6 hours later the medium is removed and is replaced with SF900 II medium minus methionine and cysteine (available from Life Technologies Inc., Rockville, MD). After 42 hours, 5 μ Ci of 35 S-methionine and 5 μ Ci 35 S-cysteine (available from Amersham) are added. The cells are further incubated for 16 hours and then are harvested by centrifugation. The proteins in the supernatant as well as the intracellular proteins are analyzed by SDS-PAGE followed by autoradiography (if radiolabeled).

Microsequencing of the amino acid sequence of the amino terminus of purified protein may be used to determine the amino terminal sequence of the produced protein.

Example 8: Expression of a Polypeptide in Mammalian Cells

The polypeptide of the present invention can be expressed in a mammalian cell. A typical mammalian expression vector contains a promoter element, which mediates the initiation of transcription of mRNA, a protein coding sequence, and signals required for the termination of transcription and polyadenylation of the transcript. Additional elements include enhancers, Kozak sequences and intervening sequences flanked by donor and acceptor sites for RNA splicing. Highly efficient transcription is achieved with the early and late promoters from SV40, the long terminal repeats (LTRs) from Retroviruses, e.g., RSV, HTLV, HIV and the early promoter of the cytomegalovirus (CMV). However, cellular elements can also be used (e.g., the human actin promoter).

Suitable expression vectors for use in practicing the present invention include, for example, vectors such as pSVL and pMSG (Pharmacia, Uppsala, Sweden), pRSVcat (ATCC 37152), pSV2dhfr (ATCC 37146), pBC12MI (ATCC 67109), pCMVSPORT 2.0, and pCMVSPORT 3.0. Mammalian host cells that could be used include, human Hela, 293, H9 and Jurkat cells, mouse NIH3T3 and C127 cells, Cos 1, Cos 7 and CV1, quail QC1-3 cells, mouse L cells and Chinese hamster ovary (CHO) cells.

Alternatively, the polypeptide can be expressed in stable cell lines containing the polynucleotide integrated into a chromosome. The co-transfection with a selectable marker such as dhfr, gpt, neomycin, hygromycin allows the identification and isolation of the transfected cells.

5 The transfected gene can also be amplified to express large amounts of the encoded protein. The DHFR (dihydrofolate reductase) marker is useful in developing cell lines that carry several hundred or even several thousand copies of the gene of interest. (See, e.g., Alt, F. W., et al., *J. Biol. Chem.* 253:1357-1370 (1978); Hamlin, J. L. and Ma, C., *Biochem. et Biophys. Acta*, 1097:107-143 (1990); Page, M. J. and
10 Sydenham, M. A., *Biotechnology* 9:64-68 (1991).) Another useful selection marker is the enzyme glutamine synthase (GS) (Murphy et al., *Biochem J.* 227:277-279 (1991); Bebbington et al., *Bio/Technology* 10:169-175 (1992). Using these markers, the mammalian cells are grown in selective medium and the cells with the highest resistance are selected. These cell lines contain the amplified gene(s) integrated into a
15 chromosome. Chinese hamster ovary (CHO) and NSO cells are often used for the production of proteins.

Derivatives of the plasmid pSV2-dhfr (ATCC Accession No. 37146), the expression vectors pC4 (ATCC Accession No. 209646) and pC6 (ATCC Accession No. 209647) contain the strong promoter (LTR) of the Rous Sarcoma Virus (Cullen et al., *Molecular and Cellular Biology*, 438-447 (March, 1985)) plus a fragment of the
20 CMV-enhancer (Boshart et al., *Cell* 41:521-530 (1985).) Multiple cloning sites, e.g., with the restriction enzyme cleavage sites BamHI, XbaI and Asp718, facilitate the cloning of the gene of interest. The vectors also contain the 3' intron, the polyadenylation and termination signal of the rat preproinsulin gene, and the mouse
25 DHFR gene under control of the SV40 early promoter.

Specifically, the plasmid pC6, for example, is digested with appropriate restriction enzymes and then dephosphorylated using calf intestinal phosphates by procedures known in the art. The vector is then isolated from a 1% agarose gel.

A polynucleotide of the present invention is amplified according to the protocol
30 outlined in Example 1. If the naturally occurring signal sequence is used to produce the secreted protein, the vector does not need a second signal peptide. Alternatively, if the naturally occurring signal sequence is not used, the vector can be modified to include a heterologous signal sequence. (See, e.g., WO 96/34891.)

The amplified fragment is isolated from a 1% agarose gel using a commercially
35 available kit ("GeneClean," BIO 101 Inc., La Jolla, Ca.). The fragment then is digested with appropriate restriction enzymes and again purified on a 1% agarose gel.

The amplified fragment is then digested with the same restriction enzyme and purified on a 1% agarose gel. The isolated fragment and the dephosphorylated vector are then ligated with T4 DNA ligase. *E. coli* HB101 or XL-1 Blue cells are then transformed and bacteria are identified that contain the fragment inserted into plasmid pC6 using, for instance, restriction enzyme analysis.

Chinese hamster ovary cells lacking an active DHFR gene is used for transfection. Five μ g of the expression plasmid pC6 is cotransfected with 0.5 μ g of the plasmid pSVneo using lipofectin (Felgner et al., *supra*). The plasmid pSV2-neo contains a dominant selectable marker, the *neo* gene from Tn5 encoding an enzyme that confers resistance to a group of antibiotics including G418. The cells are seeded in alpha minus MEM supplemented with 1 mg/ml G418. After 2 days, the cells are trypsinized and seeded in hybridoma cloning plates (Greiner, Germany) in alpha minus MEM supplemented with 10, 25, or 50 ng/ml of methotrexate plus 1 mg/ml G418. After about 10-14 days single clones are trypsinized and then seeded in 6-well petri dishes or 10 ml flasks using different concentrations of methotrexate (50 nM, 100 nM, 200 nM, 400 nM, 800 nM). Clones growing at the highest concentrations of methotrexate are then transferred to new 6-well plates containing even higher concentrations of methotrexate (1 μ M, 2 μ M, 5 μ M, 10 mM, 20 mM). The same procedure is repeated until clones are obtained which grow at a concentration of 100 - 200 μ M. Expression of the desired gene product is analyzed, for instance, by SDS-PAGE and Western blot or by reversed phase HPLC analysis.

Example 9: Protein Fusions

The polypeptides of the present invention are preferably fused to other proteins. These fusion proteins can be used for a variety of applications. For example, fusion of the present polypeptides to His-tag, HA-tag, protein A, IgG domains, and maltose binding protein facilitates purification. (See Example 5; see also EP A 394,827; Traunecker, et al., Nature 331:84-86 (1988).) Similarly, fusion to IgG-1, IgG-3, and albumin increases the halflife time in vivo. Nuclear localization signals fused to the polypeptides of the present invention can target the protein to a specific subcellular localization, while covalent heterodimer or homodimers can increase or decrease the activity of a fusion protein. Fusion proteins can also create chimeric molecules having more than one function. Finally, fusion proteins can increase solubility and/or stability of the fused protein compared to the non-fused protein. All of the types of fusion proteins described above can be made by modifying the following protocol, which outlines the fusion of a polypeptide to an IgG molecule, or the protocol described in Example 5.

Briefly, the human Fc portion of the IgG molecule can be PCR amplified, using primers that span the 5' and 3' ends of the sequence described below. These primers also should have convenient restriction enzyme sites that will facilitate cloning into an expression vector, preferably a mammalian expression vector.

5 For example, if pC4 (Accession No. 209646) is used, the human Fc portion can be ligated into the BamHI cloning site. Note that the 3' BamHI site should be destroyed. Next, the vector containing the human Fc portion is re-restricted with BamHI, linearizing the vector, and a polynucleotide of the present invention, isolated by the PCR protocol described in Example 1, is ligated into this BamHI site. Note that
10 the polynucleotide is cloned without a stop codon, otherwise a fusion protein will not be produced.

If the naturally occurring signal sequence is used to produce the secreted protein, pC4 does not need a second signal peptide. Alternatively, if the naturally occurring signal sequence is not used, the vector can be modified to include a
15 heterologous signal sequence. (See, e.g., WO 96/34891.)

Human IgG Fc region:

GGGATCCGGAGCCCAAATCTTCTGACAAAACCTCACACATGCCCACCGTGCC
CAGCACCTGAATTCGAGGGTGCACCGTCAGTCTTCCTCTTCCCCCAAAC
20 CAAGGACACCCTCATGATCTCCCGGACTCCTGAGGTCACATGCGTGGTGGT
GGACGTAAGCCACGAAGACCCTGAGGTCAAGTTCAACTGGTACGTGGACG
GCGTGGAGGTGCATAATGCCAAGACAAAGCCGCGGGAGGAGCAGTACAAC
AGCACGTACCGTGTGGTCAGCGTCCTCACCGTCCTGCACCAGGACTGGCTG
AATGGCAAGGAGTACAAGTGCAAGGTCTCCAACAAAGCCCTCCCAACCCCC
25 ATCGAGAAAACCATCTCCAAAGCCAAAGGGCAGCCCCGAGAACCACAGGT
GTACACCCTGCCCCCATCCCGGGATGAGCTGACCAAGAACCAGGTGAGCCT
GACCTGCCTGGTCAAAGGCTTCTATCCAAGCGACATCGCCGTGGAGTGGGA
GAGCAATGGGCAGCCGGAGAACAACACTACAAGACCACGCCTCCCGTGCTGG
ACTCCGACGGCTCCTTCTTCCTCTACAGCAAGCTCACCGTGGACAAGAGCA
30 GGTGGCAGCAGGGGAACGTCTTCTCATGCTCCGTGATGCATGAGGCTCTGC
ACAACCACTACACGCAGAAGAGCCTCTCCCTGTCTCCGGGTAAATGAGTGC
GACGGCCGCGACTCTAGAGGAT (SEQ ID NO:1)

Example 10: Production of an Antibody from a Polypeptide

35 The antibodies of the present invention can be prepared by a variety of methods. (See, Current Protocols, Chapter 2.) For example, cells expressing a polypeptide of the present invention is administered to an animal to induce the production of sera

containing polyclonal antibodies. In a preferred method, a preparation of the secreted protein is prepared and purified to render it substantially free of natural contaminants. Such a preparation is then introduced into an animal in order to produce polyclonal antisera of greater specific activity.

5 In the most preferred method, the antibodies of the present invention are monoclonal antibodies (or protein binding fragments thereof). Such monoclonal antibodies can be prepared using hybridoma technology. (Köhler et al., *Nature* 256:495 (1975); Köhler et al., *Eur. J. Immunol.* 6:511 (1976); Köhler et al., *Eur. J. Immunol.* 6:292 (1976); Hammerling et al., in: *Monoclonal Antibodies and T-Cell*
10 *Hybridomas*, Elsevier, N.Y., pp. 563-681 (1981).) In general, such procedures involve immunizing an animal (preferably a mouse) with polypeptide or, more preferably, with a secreted polypeptide-expressing cell. Such cells may be cultured in any suitable tissue culture medium; however, it is preferable to culture cells in Earle's modified Eagle's medium supplemented with 10% fetal bovine serum (inactivated at
15 about 56°C), and supplemented with about 10 g/l of nonessential amino acids, about 1,000 U/ml of penicillin, and about 100 µg/ml of streptomycin.

The splenocytes of such mice are extracted and fused with a suitable myeloma cell line. Any suitable myeloma cell line may be employed in accordance with the present invention; however, it is preferable to employ the parent myeloma cell line
20 (SP2O), available from the ATCC. After fusion, the resulting hybridoma cells are selectively maintained in HAT medium, and then cloned by limiting dilution as described by Wands et al. (*Gastroenterology* 80:225-232 (1981).) The hybridoma cells obtained through such a selection are then assayed to identify clones which secrete antibodies capable of binding the polypeptide.

25 Alternatively, additional antibodies capable of binding to the polypeptide can be produced in a two-step procedure using anti-idiotypic antibodies. Such a method makes use of the fact that antibodies are themselves antigens, and therefore, it is possible to obtain an antibody which binds to a second antibody. In accordance with this method, protein specific antibodies are used to immunize an animal, preferably a
30 mouse. The splenocytes of such an animal are then used to produce hybridoma cells, and the hybridoma cells are screened to identify clones which produce an antibody whose ability to bind to the protein-specific antibody can be blocked by the polypeptide. Such antibodies comprise anti-idiotypic antibodies to the protein-specific antibody and can be used to immunize an animal to induce formation of further protein-specific
35 antibodies.

It will be appreciated that Fab and F(ab')₂ and other fragments of the antibodies of the present invention may be used according to the methods disclosed herein. Such fragments are typically produced by proteolytic cleavage, using enzymes such as papain (to produce Fab fragments) or pepsin (to produce F(ab')₂ fragments). Alternatively, secreted protein-binding fragments can be produced through the application of recombinant DNA technology or through synthetic chemistry.

For in vivo use of antibodies in humans, it may be preferable to use "humanized" chimeric monoclonal antibodies. Such antibodies can be produced using genetic constructs derived from hybridoma cells producing the monoclonal antibodies described above. Methods for producing chimeric antibodies are known in the art. (See, for review, Morrison, Science 229:1202 (1985); Oi et al., BioTechniques 4:214 (1986); Cabilly et al., U.S. Patent No. 4,816,567; Taniguchi et al., EP 171496; Morrison et al., EP 173494; Neuberger et al., WO 8601533; Robinson et al., WO 8702671; Boulianne et al., Nature 312:643 (1984); Neuberger et al., Nature 314:268 (1985).)

Example 11: Production Of Secreted Protein For High-Throughput Screening Assays

The following protocol produces a supernatant containing a polypeptide to be tested. This supernatant can then be used in the Screening Assays described in Examples 13-20.

First, dilute Poly-D-Lysine (644 587 Boehringer-Mannheim) stock solution (1mg/ml in PBS) 1:20 in PBS (w/o calcium or magnesium 17-516F Biowhittaker) for a working solution of 50ug/ml. Add 200 ul of this solution to each well (24 well plates) and incubate at RT for 20 minutes. Be sure to distribute the solution over each well (note: a 12-channel pipetter may be used with tips on every other channel). Aspirate off the Poly-D-Lysine solution and rinse with 1ml PBS (Phosphate Buffered Saline). The PBS should remain in the well until just prior to plating the cells and plates may be poly-lysine coated in advance for up to two weeks.

Plate 293T cells (do not carry cells past P+20) at 2×10^5 cells/well in .5ml DMEM(Dulbecco's Modified Eagle Medium)(with 4.5 G/L glucose and L-glutamine (12-604F Biowhittaker))/10% heat inactivated FBS(14-503F Biowhittaker)/1x Penstrep(17-602E Biowhittaker). Let the cells grow overnight.

The next day, mix together in a sterile solution basin: 300 ul Lipofectamine (18324-012 Gibco/BRL) and 5ml Optimem I (31985070 Gibco/BRL)/96-well plate. With a small volume multi-channel pipetter, aliquot approximately 2ug of an expression vector containing a polynucleotide insert, produced by the methods described in

- Examples 8 or 9, into an appropriately labeled 96-well round bottom plate. With a multi-channel pipetter, add 50ul of the Lipofectamine/Optimem I mixture to each well. Pipette up and down gently to mix. Incubate at RT 15-45 minutes. After about 20 minutes, use a multi-channel pipetter to add 150ul Optimem I to each well. As a control, one plate of vector DNA lacking an insert should be transfected with each set of transfections.

- Preferably, the transfection should be performed by tag-teaming the following tasks. By tag-teaming, hands on time is cut in half, and the cells do not spend too much time on PBS. First, person A aspirates off the media from four 24-well plates of cells, and then person B rinses each well with .5-1ml PBS. Person A then aspirates off PBS rinse, and person B, using a 12-channel pipetter with tips on every other channel, adds the 200ul of DNA/Lipofectamine/Optimem I complex to the odd wells first, then to the even wells, to each row on the 24-well plates. Incubate at 37°C for 6 hours.

- While cells are incubating, prepare appropriate media, either 1% BSA in DMEM with 1x penstrep, or CHO-5 media (116.6 mg/L of CaCl₂ (anhyd); 0.00130 mg/L CuSO₄·5H₂O; 0.050 mg/L of Fe(NO₃)₃·9H₂O; 0.417 mg/L of FeSO₄·7H₂O; 311.80 mg/L of KCl; 28.64 mg/L of MgCl₂; 48.84 mg/L of MgSO₄; 6995.50 mg/L of NaCl; 2400.0 mg/L of NaHCO₃; 62.50 mg/L of NaH₂PO₄·H₂O; 71.02 mg/L of Na₂HPO₄; .4320 mg/L of ZnSO₄·7H₂O; .002 mg/L of Arachidonic Acid; 1.022 mg/L of Cholesterol; .070 mg/L of DL-alpha-Tocopherol-Acetate; 0.0520 mg/L of Linoleic Acid; 0.010 mg/L of Linolenic Acid; 0.010 mg/L of Myristic Acid; 0.010 mg/L of Oleic Acid; 0.010 mg/L of Palmitic Acid; 0.010 mg/L of Palmitic Acid; 100 mg/L of Pluronic F-68; 0.010 mg/L of Stearic Acid; 2.20 mg/L of Tween 80; 4551 mg/L of D-Glucose; 130.85 mg/ml of L-Alanine; 147.50 mg/ml of L-Arginine-HCL; 7.50 mg/ml of L-Asparagine-H₂O; 6.65 mg/ml of L-Aspartic Acid; 29.56 mg/ml of L-Cystine-2HCL-H₂O; 31.29 mg/ml of L-Cystine-2HCL; 7.35 mg/ml of L-Glutamic Acid; 365.0 mg/ml of L-Glutamine; 18.75 mg/ml of Glycine; 52.48 mg/ml of L-Histidine-HCL-H₂O; 106.97 mg/ml of L-Isoleucine; 111.45 mg/ml of L-Leucine; 163.75 mg/ml of L-Lysine HCL; 32.34 mg/ml of L-Methionine; 68.48 mg/ml of L-Phenylalanine; 40.0 mg/ml of L-Proline; 26.25 mg/ml of L-Serine; 101.05 mg/ml of L-Threonine; 19.22 mg/ml of L-Tryptophan; 91.79 mg/ml of L-Tyrosine-2Na-2H₂O; 99.65 mg/ml of L-Valine; 0.0035 mg/L of Biotin; 3.24 mg/L of D-Ca Pantothenate; 11.78 mg/L of Choline Chloride; 4.65 mg/L of Folic Acid; 15.60 mg/L of i-Inositol; 3.02 mg/L of Niacinamide; 3.00 mg/L of Pyridoxal HCL; 0.031 mg/L of Pyridoxine HCL; 0.319 mg/L of Riboflavin; 3.17 mg/L of Thiamine HCL; 0.365 mg/L of Thymidine; and 0.680 mg/L of Vitamin B₁₂; 25 mM of HEPES Buffer; 2.39 mg/L of Na Hypoxanthine;

0.105 mg/L of Lipoic Acid; 0.081 mg/L of Sodium Putrescine-2HCL; 55.0 mg/L of Sodium Pyruvate; 0.0067 mg/L of Sodium Selenite; 20uM of Ethanolamine; 0.122 mg/L of Ferric Citrate; 41.70 mg/L of Methyl-B-Cyclodextrin complexed with Linoleic Acid; 33.33 mg/L of Methyl-B-Cyclodextrin complexed with Oleic Acid; and 10 mg/L of Methyl-B-Cyclodextrin complexed with Retinal) with 2mm glutamine and 1x penstrep. (BSA (81-068-3 Bayer) 100gm dissolved in 1L DMEM for a 10% BSA stock solution). Filter the media and collect 50 ul for endotoxin assay in 15ml polystyrene conical.

The transfection reaction is terminated, preferably by tag-teaming, at the end of the incubation period. Person A aspirates off the transfection media, while person B adds 1.5ml appropriate media to each well. Incubate at 37°C for 45 or 72 hours depending on the media used: 1%BSA for 45 hours or CHO-5 for 72 hours.

On day four, using a 300ul multichannel pipetter, aliquot 600ul in one 1ml deep well plate and the remaining supernatant into a 2ml deep well. The supernatants from each well can then be used in the assays described in Examples 13-20.

It is specifically understood that when activity is obtained in any of the assays described below using a supernatant, the activity originates from either the polypeptide directly (e.g., as a secreted protein) or by the polypeptide inducing expression of other proteins, which are then secreted into the supernatant. Thus, the invention further provides a method of identifying the protein in the supernatant characterized by an activity in a particular assay.

Example 12: Construction of GAS Reporter Construct

One signal transduction pathway involved in the differentiation and proliferation of cells is called the Jaks-STATs pathway. Activated proteins in the Jaks-STATs pathway bind to gamma activation site "GAS" elements or interferon-sensitive responsive element ("ISRE"), located in the promoter of many genes. The binding of a protein to these elements alter the expression of the associated gene.

GAS and ISRE elements are recognized by a class of transcription factors called Signal Transducers and Activators of Transcription, or "STATs." There are six members of the STATs family. Stat1 and Stat3 are present in many cell types, as is Stat2 (as response to IFN-alpha is widespread). Stat4 is more restricted and is not in many cell types though it has been found in T helper class I, cells after treatment with IL-12. Stat5 was originally called mammary growth factor, but has been found at higher concentrations in other cells including myeloid cells. It can be activated in tissue culture cells by many cytokines.

The STATs are activated to translocate from the cytoplasm to the nucleus upon tyrosine phosphorylation by a set of kinases known as the Janus Kinase ("Jaks") family. Jaks represent a distinct family of soluble tyrosine kinases and include Tyk2, Jak1, Jak2, and Jak3. These kinases display significant sequence similarity and are generally catalytically inactive in resting cells.

The Jaks are activated by a wide range of receptors summarized in the Table below. (Adapted from review by Schidler and Darnell, *Ann. Rev. Biochem.* 64:621-51 (1995).) A cytokine receptor family, capable of activating Jaks, is divided into two groups: (a) Class 1 includes receptors for IL-2, IL-3, IL-4, IL-6, IL-7, IL-9, IL-11, IL-12, IL-15, Epo, PRL, GH, G-CSF, GM-CSF, LIF, CNTF, and thrombopoietin; and (b) Class 2 includes IFN- α , IFN- γ , and IL-10. The Class 1 receptors share a conserved cysteine motif (a set of four conserved cysteines and one tryptophan) and a WSXWS motif (a membrane proximal region encoding Trp-Ser-Xxx-Trp-Ser (SEQ ID NO:2)).

Thus, on binding of a ligand to a receptor, Jaks are activated, which in turn activate STATs, which then translocate and bind to GAS elements. This entire process is encompassed in the Jaks-STATs signal transduction pathway.

Therefore, activation of the Jaks-STATs pathway, reflected by the binding of the GAS or the ISRE element, can be used to indicate proteins involved in the proliferation and differentiation of cells. For example, growth factors and cytokines are known to activate the Jaks-STATs pathway. (See Table below.) Thus, by using GAS elements linked to reporter molecules, activators of the Jaks-STATs pathway can be identified.

	<u>Ligand</u>	<u>tyk2</u>	<u>JAKs</u> <u>Jak1</u>	<u>Jak2</u>	<u>Jak3</u>	<u>STATS</u>	<u>GAS(elements) or ISRE</u>
	<u>IFN family</u>						
5	IFN-a/B	+	+	-	-	1,2,3	ISRE
	IFN-g		+	+	-	1	GAS (IRF1>Lys6>IFP)
	Il-10	+	?	?	-	1,3	
	<u>gp130 family</u>						
10	IL-6 (Pleiotrohic)	+	+	+	?	1,3	GAS (IRF1>Lys6>IFP)
	Il-11(Pleiotrohic)	?	+	?	?	1,3	
	OnM(Pleiotrohic)	?	+	+	?	1,3	
	LIF(Pleiotrohic)	?	+	+	?	1,3	
	CNTF(Pleiotrohic)	-/+	+	+	?	1,3	
15	G-CSF(Pleiotrohic)	?	+	?	?	1,3	
	IL-12(Pleiotrohic)	+	-	+	+	1,3	
	<u>g-C family</u>						
	IL-2 (lymphocytes)	-	+	-	+	1,3,5	GAS
20	IL-4 (lymph/myeloid)	-	+	-	+	6	GAS (IRF1 = IFP >>Ly6)(IgH)
	IL-7 (lymphocytes)	-	+	-	+	5	GAS
	IL-9 (lymphocytes)	-	+	-	+	5	GAS
	IL-13 (lymphocyte)	-	+	?	?	6	GAS
	IL-15	?	+	?	+	5	GAS
25	<u>gp140 family</u>						
	IL-3 (myeloid)	-	-	+	-	5	GAS (IRF1>IFP>>Ly6)
	IL-5 (myeloid)	-	-	+	-	5	GAS
30	GM-CSF (myeloid)	-	-	+	-	5	GAS
	<u>Growth hormone family</u>						
	GH	?	-	+	-	5	
	PRL	?	+/-	+	-	1,3,5	
35	EPO	?	-	+	-	5	GAS(B-CAS>IRF1=IFP>>Ly6)
	<u>Receptor Tyrosine Kinases</u>						
	EGF	?	+	+	-	1,3	GAS (IRF1)
	PDGF	?	+	+	-	1,3	
40	CSF-1	?	+	+	-	1,3	GAS (not IRF1)

To construct a synthetic GAS containing promoter element, which is used in the Biological Assays described in Examples 13-14, a PCR based strategy is employed to generate a GAS-SV40 promoter sequence. The 5' primer contains four tandem copies of the GAS binding site found in the IRF1 promoter and previously demonstrated to bind STATs upon induction with a range of cytokines (Rothman et al., Immunity 1:457-468 (1994).), although other GAS or ISRE elements can be used instead. The 5' primer also contains 18bp of sequence complementary to the SV40 early promoter sequence and is flanked with an XhoI site. The sequence of the 5' primer is:
5':GCGCCTCGAGATTTCCCCGAAATCTAGATTTCCCCGAAATGATTTCCCCG
10 AAATGATTTCCCCGAAATATCTGCCATCTCAATTAG:3' (SEQ ID NO:3)

The downstream primer is complementary to the SV40 promoter and is flanked with a Hind III site: 5':GCGGCAAGCTTTTTGCAAAGCCTAGGC:3' (SEQ ID NO:4)

PCR amplification is performed using the SV40 promoter template present in the B-gal:promoter plasmid obtained from Clontech. The resulting PCR fragment is digested with XhoI/Hind III and subcloned into BLSK2-. (Stratagene.) Sequencing with forward and reverse primers confirms that the insert contains the following sequence:
5':CTCGAGATTTCCCCGAAATCTAGATTTCCCCGAAATGATTTCCCCGAAATG
20 ATTTCCCCGAAATATCTGCCATCTCAATTAGTCAGCAACCATAGTCCCCGCC
CTAACTCCGCCCATCCCGCCCCTAACTCCGCCCAGTTCGCCCATTCTCCGC
CCCATGGCTGACTAATTTTTTTTATTTATGCAGAGGCCGAGGCCGCCTCGGC
CTCTGAGCTATTCCAGAAGTAGTGAGGAGGCTTTTTTGGAGGCCTAGGCTTT
TGCAAAAAGCTT:3' (SEQ ID NO:5)

25 With this GAS promoter element linked to the SV40 promoter, a GAS:SEAP2 reporter construct is next engineered. Here, the reporter molecule is a secreted alkaline phosphatase, or "SEAP." Clearly, however, any reporter molecule can be instead of SEAP, in this or in any of the other Examples. Well known reporter molecules that can be used instead of SEAP include chloramphenicol acetyltransferase (CAT), luciferase,
30 alkaline phosphatase, B-galactosidase, green fluorescent protein (GFP), or any protein detectable by an antibody.

The above sequence confirmed synthetic GAS-SV40 promoter element is subcloned into the pSEAP-Promoter vector obtained from Clontech using HindIII and XhoI, effectively replacing the SV40 promoter with the amplified GAS:SV40 promoter element, to create the GAS-SEAP vector. However, this vector does not contain a
35 neomycin resistance gene, and therefore, is not preferred for mammalian expression systems.

Thus, in order to generate mammalian stable cell lines expressing the GAS-SEAP reporter, the GAS-SEAP cassette is removed from the GAS-SEAP vector using SalI and NotI, and inserted into a backbone vector containing the neomycin resistance gene, such as pGFP-1 (Clontech), using these restriction sites in the multiple cloning site, to create the GAS-SEAP/Neo vector. Once this vector is transfected into mammalian cells, this vector can then be used as a reporter molecule for GAS binding as described in Examples 13-14.

Other constructs can be made using the above description and replacing GAS with a different promoter sequence. For example, construction of reporter molecules containing NFK-B and EGR promoter sequences are described in Examples 15 and 16. However, many other promoters can be substituted using the protocols described in these Examples. For instance, SRE, IL-2, NFAT, or Osteocalcin promoters can be substituted, alone or in combination (e.g., GAS/NF-KB/EGR, GAS/NF-KB, IL-2/NFAT, or NF-KB/GAS). Similarly, other cell lines can be used to test reporter construct activity, such as HELA (epithelial), HUVEC (endothelial), Reh (B-cell), Saos-2 (osteoblast), HUVAC (aortic), or Cardiomyocyte.

Example 13: High-Throughput Screening Assay for T-cell Activity.

The following protocol is used to assess T-cell activity by identifying factors, such as growth factors and cytokines, that may proliferate or differentiate T-cells. T-cell activity is assessed using the GAS/SEAP/Neo construct produced in Example 12. Thus, factors that increase SEAP activity indicate the ability to activate the Jaks-STATS signal transduction pathway. The T-cell used in this assay is Jurkat T-cells (ATCC Accession No. TIB-152), although Molt-3 cells (ATCC Accession No. CRL-1552) and Molt-4 cells (ATCC Accession No. CRL-1582) cells can also be used.

Jurkat T-cells are lymphoblastic CD4+ Th1 helper cells. In order to generate stable cell lines, approximately 2 million Jurkat cells are transfected with the GAS-SEAP/neo vector using DMRIE-C (Life Technologies)(transfection procedure described below). The transfected cells are seeded to a density of approximately 20,000 cells per well and transfectants resistant to 1 mg/ml gentamicin selected. Resistant colonies are expanded and then tested for their response to increasing concentrations of interferon gamma. The dose response of a selected clone is demonstrated.

Specifically, the following protocol will yield sufficient cells for 75 wells containing 200 ul of cells. Thus, it is either scaled up, or performed in multiple to generate sufficient cells for multiple 96 well plates. Jurkat cells are maintained in RPMI + 10% serum with 1%Pen-Strep. Combine 2.5 mls of OPTI-MEM (Life Technologies)

with 10 ug of plasmid DNA in a T25 flask. Add 2.5 ml OPTI-MEM containing 50 ul of DMRIE-C and incubate at room temperature for 15-45 mins.

- During the incubation period, count cell concentration, spin down the required number of cells (10^7 per transfection), and resuspend in OPTI-MEM to a final
5 concentration of 10^7 cells/ml. Then add 1ml of 1×10^7 cells in OPTI-MEM to T25 flask and incubate at 37°C for 6 hrs. After the incubation, add 10 ml of RPMI + 15% serum.

The Jurkat:GAS-SEAP stable reporter lines are maintained in RPMI + 10% serum, 1 mg/ml Gentamicin, and 1% Pen-Strep. These cells are treated with supernatants containing a polypeptide as produced by the protocol described in Example 11.

- 10 On the day of treatment with the supernatant, the cells should be washed and resuspended in fresh RPMI + 10% serum to a density of 500,000 cells per ml. The exact number of cells required will depend on the number of supernatants being screened. For one 96 well plate, approximately 10 million cells (for 10 plates, 100 million cells) are required.

- 15 Transfer the cells to a triangular reservoir boat, in order to dispense the cells into a 96 well dish, using a 12 channel pipette. Using a 12 channel pipette, transfer 200 ul of cells into each well (therefore adding 100, 000 cells per well).

- After all the plates have been seeded, 50 ul of the supernatants are transferred directly from the 96 well plate containing the supernatants into each well using a 12
20 channel pipette. In addition, a dose of exogenous interferon gamma (0.1, 1.0, 10 ng) is added to wells H9, H10, and H11 to serve as additional positive controls for the assay.

- The 96 well dishes containing Jurkat cells treated with supernatants are placed in an incubator for 48 hrs (note: this time is variable between 48-72 hrs). 35 ul samples
25 from each well are then transferred to an opaque 96 well plate using a 12 channel pipette. The opaque plates should be covered (using sellophane covers) and stored at -20°C until SEAP assays are performed according to Example 17. The plates containing the remaining treated cells are placed at 4°C and serve as a source of material for repeating the assay on a specific well if desired.

- 30 As a positive control, 100 Unit/ml interferon gamma can be used which is known to activate Jurkat T cells. Over 30 fold induction is typically observed in the positive control wells.

35 **Example 14: High-Throughput Screening Assay Identifying Myeloid Activity**

The following protocol is used to assess myeloid activity by identifying factors, such as growth factors and cytokines, that may proliferate or differentiate myeloid cells. Myeloid cell activity is assessed using the GAS/SEAP/Neo construct produced in Example 12. Thus, factors that increase SEAP activity indicate the ability to activate the Jaks-STATS signal transduction pathway. The myeloid cell used in this assay is U937, a pre-monocyte cell line, although TF-1, HL60, or KG1 can be used.

To transiently transfect U937 cells with the GAS/SEAP/Neo construct produced in Example 12, a DEAE-Dextran method (Kharbanda et. al., 1994, Cell Growth & Differentiation, 5:259-265) is used. First, harvest 2×10^7 U937 cells and wash with PBS. The U937 cells are usually grown in RPMI 1640 medium containing 10% heat-inactivated fetal bovine serum (FBS) supplemented with 100 units/ml penicillin and 100 mg/ml streptomycin.

Next, suspend the cells in 1 ml of 20 mM Tris-HCl (pH 7.4) buffer containing 0.5 mg/ml DEAE-Dextran, 8 ug GAS-SEAP2 plasmid DNA, 140 mM NaCl, 5 mM KCl, 375 uM $\text{Na}_2\text{HPO}_4 \cdot 7\text{H}_2\text{O}$, 1 mM MgCl_2 , and 675 uM CaCl_2 . Incubate at 37°C for 45 min.

Wash the cells with RPMI 1640 medium containing 10% FBS and then resuspend in 10 ml complete medium and incubate at 37°C for 36 hr.

The GAS-SEAP/U937 stable cells are obtained by growing the cells in 400 ug/ml G418. The G418-free medium is used for routine growth but every one to two months, the cells should be re-grown in 400 ug/ml G418 for couple of passages.

These cells are tested by harvesting 1×10^8 cells (this is enough for ten 96-well plates assay) and wash with PBS. Suspend the cells in 200 ml above described growth medium, with a final density of 5×10^5 cells/ml. Plate 200 ul cells per well in the 96-well plate (or 1×10^5 cells/well).

Add 50 ul of the supernatant prepared by the protocol described in Example 11. Incubate at 37°C for 48 to 72 hr. As a positive control, 100 Unit/ml interferon gamma can be used which is known to activate U937 cells. Over 30 fold induction is typically observed in the positive control wells. SEAP assay the supernatant according to the protocol described in Example 17.

Example 15: High-Throughput Screening Assay Identifying Neuronal Activity.

When cells undergo differentiation and proliferation, a group of genes are activated through many different signal transduction pathways. One of these genes, EGR1 (early growth response gene 1), is induced in various tissues and cell types upon

activation. The promoter of EGR1 is responsible for such induction. Using the EGR1 promoter linked to reporter molecules, activation of cells can be assessed.

Particularly, the following protocol is used to assess neuronal activity in PC12 cell lines. PC12 cells (rat phenochromocytoma cells) are known to proliferate and/or
5 differentiate by activation with a number of mitogens, such as TPA (tetradecanoyl phorbol acetate), NGF (nerve growth factor), and EGF (epidermal growth factor). The EGR1 gene expression is activated during this treatment. Thus, by stably transfecting PC12 cells with a construct containing an EGR promoter linked to SEAP reporter, activation of PC12 cells can be assessed.

10 The EGR/SEAP reporter construct can be assembled by the following protocol. The EGR-1 promoter sequence (-633 to +1)(Sakamoto K et al., Oncogene 6:867-871 (1991)) can be PCR amplified from human genomic DNA using the following primers:

5' GCGCTCGAGGGATGACAGCGATAGAACCCCGG -3' (SEQ ID NO:6)

5' GCGAAGCTTCGCGACTCCCCGGATCCGCCTC-3' (SEQ ID NO:7)

15 Using the GAS:SEAP/Neo vector produced in Example 12, EGR1 amplified product can then be inserted into this vector. Linearize the GAS:SEAP/Neo vector using restriction enzymes XhoI/HindIII, removing the GAS/SV40 stuffer. Restrict the EGR1 amplified product with these same enzymes. Ligate the vector and the EGR1 promoter.

20 To prepare 96 well-plates for cell culture, two mls of a coating solution (1:30 dilution of collagen type I (Upstate Biotech Inc. Cat#08-115) in 30% ethanol (filter sterilized)) is added per one 10 cm plate or 50 ml per well of the 96-well plate, and allowed to air dry for 2 hr.

PC12 cells are routinely grown in RPMI-1640 medium (Bio Whittaker)
25 containing 10% horse serum (JRH BIOSCIENCES, Cat. # 12449-78P), 5% heat-inactivated fetal bovine serum (FBS) supplemented with 100 units/ml penicillin and 100 ug/ml streptomycin on a precoated 10 cm tissue culture dish. One to four split is done every three to four days. Cells are removed from the plates by scraping and resuspended with pipetting up and down for more than 15 times.

30 Transfect the EGR/SEAP/Neo construct into PC12 using the Lipofectamine protocol described in Example 11. EGR-SEAP/PC12 stable cells are obtained by growing the cells in 300 ug/ml G418. The G418-free medium is used for routine growth but every one to two months, the cells should be re-grown in 300 ug/ml G418 for couple of passages.

35 To assay for neuronal activity, a 10 cm plate with cells around 70 to 80% confluent is screened by removing the old medium. Wash the cells once with PBS

(Phosphate buffered saline). Then starve the cells in low serum medium (RPMI-1640 containing 1% horse serum and 0.5% FBS with antibiotics) overnight.

The next morning, remove the medium and wash the cells with PBS. Scrape off the cells from the plate, suspend the cells well in 2 ml low serum medium. Count
5 the cell number and add more low serum medium to reach final cell density as 5×10^5 cells/ml.

Add 200 μ l of the cell suspension to each well of 96-well plate (equivalent to 1×10^5 cells/well). Add 50 μ l supernatant produced by Example 11, 37°C for 48 to 72 hr. As a positive control, a growth factor known to activate PC12 cells through EGR
10 can be used, such as 50 ng/ μ l of Neuronal Growth Factor (NGF). Over fifty-fold induction of SEAP is typically seen in the positive control wells. SEAP assay the supernatant according to Example 17.

Example 16: High-Throughput Screening Assay for T-cell Activity

15 NF- κ B (Nuclear Factor κ B) is a transcription factor activated by a wide variety of agents including the inflammatory cytokines IL-1 and TNF, CD30 and CD40, lymphotoxin-alpha and lymphotoxin-beta, by exposure to LPS or thrombin, and by expression of certain viral gene products. As a transcription factor, NF- κ B regulates the expression of genes involved in immune cell activation, control of apoptosis (NF-
20 κ B appears to shield cells from apoptosis), B and T-cell development, anti-viral and antimicrobial responses, and multiple stress responses.

In non-stimulated conditions, NF- κ B is retained in the cytoplasm with I- κ B (Inhibitor κ B). However, upon stimulation, I- κ B is phosphorylated and degraded, causing NF- κ B to shuttle to the nucleus, thereby activating transcription of target
25 genes. Target genes activated by NF- κ B include IL-2, IL-6, GM-CSF, ICAM-1 and class I MHC.

Due to its central role and ability to respond to a range of stimuli, reporter constructs utilizing the NF- κ B promoter element are used to screen the supernatants produced in Example 11. Activators or inhibitors of NF- κ B would be useful in treating
30 diseases. For example, inhibitors of NF- κ B could be used to treat those diseases related to the acute or chronic activation of NF- κ B, such as rheumatoid arthritis.

To construct a vector containing the NF- κ B promoter element, a PCR based strategy is employed. The upstream primer contains four tandem copies of the NF- κ B binding site (GGGGACTTTCCC) (SEQ ID NO:8), 18 bp of sequence complementary to the 5' end of the SV40 early promoter sequence, and is flanked with an XhoI site:

5 5':GCGGCCTCGAGGGGACTTTCCCGGGGACTTTCCGGGGACTTTCCGGGAC
TTTCCATCCTGCCATCTCAATTAG:3' (SEQ ID NO:9)

The downstream primer is complementary to the 3' end of the SV40 promoter and is flanked with a Hind III site:

5':GCGGCAAGCTTTTGGCAAAGCCTAGGC:3' (SEQ ID NO:4)

10 PCR amplification is performed using the SV40 promoter template present in the pB-gal:promoter plasmid obtained from Clontech. The resulting PCR fragment is digested with XhoI and Hind III and subcloned into BLSK2-. (Stratagene) Sequencing with the T7 and T3 primers confirms the insert contains the following sequence:

15 5':CTCGAGGGGACTTTCCCGGGGACTTTCCGGGGACTTTCCGGGACTTTCC
ATCTGCCATCTCAATTAGTCAGCAACCATAGTCCCGCCCCTAACTCCGCCCA
TCCCGCCCCTAACTCCGCCCAAGTTCCGCCCATTTCTCCGCCCCATGGCTGACT
AATTTTTTTTATTTATGCAGAGGCCGAGGCCCTCGGCCTCTGAGCTATTC
20 CAGAAGTAGTGAGGAGGCTTTTTTGGAGGCCTAGGCTTTTGCAAAAAGCTT:
3' (SEQ ID NO:10)

Next, replace the SV40 minimal promoter element present in the pSEAP2-promoter plasmid (Clontech) with this NF- κ B/SV40 fragment using XhoI and HindIII.

25 However, this vector does not contain a neomycin resistance gene, and therefore, is not preferred for mammalian expression systems.

In order to generate stable mammalian cell lines, the NF- κ B/SV40/SEAP cassette is removed from the above NF- κ B/SEAP vector using restriction enzymes SalI and NotI, and inserted into a vector containing neomycin resistance. Particularly, the

30 NF- κ B/SV40/SEAP cassette was inserted into pGFP-1 (Clontech), replacing the GFP gene, after restricting pGFP-1 with SalI and NotI.

Once NF- κ B/SV40/SEAP/Neo vector is created, stable Jurkat T-cells are created and maintained according to the protocol described in Example 13. Similarly, the method for assaying supernatants with these stable Jurkat T-cells is also described

in Example 13. As a positive control, exogenous TNF alpha (0.1, 1, 10 ng) is added to wells H9, H10, and H11, with a 5-10 fold activation typically observed.

Example 17: Assay for SEAP Activity

5 As a reporter molecule for the assays described in Examples 13-16, SEAP activity is assayed using the Tropix Phospho-light Kit (Cat. BP-400) according to the following general procedure. The Tropix Phospho-light Kit supplies the Dilution, Assay, and Reaction Buffers used below.

10 Prime a dispenser with the 2.5x Dilution Buffer and dispense 15 μ l of 2.5x dilution buffer into Optiplates containing 35 μ l of a supernatant. Seal the plates with a plastic sealer and incubate at 65°C for 30 min. Separate the Optiplates to avoid uneven heating.

Cool the samples to room temperature for 15 minutes. Empty the dispenser and prime with the Assay Buffer. Add 50 μ l Assay Buffer and incubate at room temperature 5 min. Empty the dispenser and prime with the Reaction Buffer (see the table below). Add 50 μ l Reaction Buffer and incubate at room temperature for 20 minutes. Since the intensity of the chemiluminescent signal is time dependent, and it takes about 10 minutes to read 5 plates on luminometer, one should treat 5 plates at each time and start the second set 10 minutes later.

20 Read the relative light unit in the luminometer. Set H12 as blank, and print the results. An increase in chemiluminescence indicates reporter activity.

Reaction Buffer Formulation:

# of plates	Rxn buffer diluent (ml)	CSPD (ml)
10	60	3
11	65	3.25
12	70	3.5
13	75	3.75
14	80	4
15	85	4.25
16	90	4.5
17	95	4.75
18	100	5
19	105	5.25
20	110	5.5
21	115	5.75
22	120	6
23	125	6.25
24	130	6.5
25	135	6.75
26	140	7
27	145	7.25

28	150	7.5
29	155	7.75
30	160	8
31	165	8.25
32	170	8.5
33	175	8.75
34	180	9
35	185	9.25
36	190	9.5
37	195	9.75
38	200	10
39	205	10.25
40	210	10.5
41	215	10.75
42	220	11
43	225	11.25
44	230	11.5
45	235	11.75
46	240	12
47	245	12.25
48	250	12.5
49	255	12.75
50	260	13

Example 18: High-Throughput Screening Assay Identifying Changes in Small Molecule Concentration and Membrane Permeability

5 Binding of a ligand to a receptor is known to alter intracellular levels of small molecules, such as calcium, potassium, sodium, and pH, as well as alter membrane potential. These alterations can be measured in an assay to identify supernatants which bind to receptors of a particular cell. Although the following protocol describes an assay for calcium, this protocol can easily be modified to detect changes in potassium, sodium, pH, membrane potential, or any other small molecule which is detectable by a
10 fluorescent probe.

The following assay uses Fluorometric Imaging Plate Reader ("FLIPR") to measure changes in fluorescent molecules (Molecular Probes) that bind small molecules. Clearly, any fluorescent molecule detecting a small molecule can be used instead of the calcium fluorescent molecule, fluo-3, used here.

15 For adherent cells, seed the cells at 10,000 -20,000 cells/well in a Co-star black 96-well plate with clear bottom. The plate is incubated in a CO₂ incubator for 20 hours. The adherent cells are washed two times in Biotek washer with 200 ul of HBSS (Hank's Balanced Salt Solution) leaving 100 ul of buffer after the final wash.

20 A stock solution of 1 mg/ml fluo-3 is made in 10% pluronic acid DMSO. To load the cells with fluo-3, 50 ul of 12 ug/ml fluo-3 is added to each well. The plate is

incubated at 37°C in a CO₂ incubator for 60 min. The plate is washed four times in the Biotek washer with HBSS leaving 100 ul of buffer.

For non-adherent cells, the cells are spun down from culture media. Cells are re-suspended to 2-5x10⁶ cells/ml with HBSS in a 50-ml conical tube. 4 ul of 1 mg/ml fluo-3 solution in 10% pluronic acid DMSO is added to each ml of cell suspension. The tube is then placed in a 37°C water bath for 30-60 min. The cells are washed twice with HBSS, resuspended to 1x10⁶ cells/ml, and dispensed into a microplate, 100 ul/well. The plate is centrifuged at 1000 rpm for 5 min. The plate is then washed once in Denley CellWash with 200 ul, followed by an aspiration step to 100 ul final volume.

For a non-cell based assay, each well contains a fluorescent molecule, such as fluo-3. The supernatant is added to the well, and a change in fluorescence is detected.

To measure the fluorescence of intracellular calcium, the FLIPR is set for the following parameters: (1) System gain is 300-800 mW; (2) Exposure time is 0.4 second; (3) Camera F/stop is F/2; (4) Excitation is 488 nm; (5) Emission is 530 nm; and (6) Sample addition is 50 ul. Increased emission at 530 nm indicates an extracellular signaling event which has resulted in an increase in the intracellular Ca⁺⁺ concentration.

Example 19: High-Throughput Screening Assay Identifying Tyrosine Kinase Activity

The Protein Tyrosine Kinases (PTK) represent a diverse group of transmembrane and cytoplasmic kinases. Within the Receptor Protein Tyrosine Kinase (RPTK) group are receptors for a range of mitogenic and metabolic growth factors including the PDGF, FGF, EGF, NGF, HGF and Insulin receptor subfamilies. In addition there are a large family of RPTKs for which the corresponding ligand is unknown. Ligands for RPTKs include mainly secreted small proteins, but also membrane-bound and extracellular matrix proteins.

Activation of RPTK by ligands involves ligand-mediated receptor dimerization, resulting in transphosphorylation of the receptor subunits and activation of the cytoplasmic tyrosine kinases. The cytoplasmic tyrosine kinases include receptor associated tyrosine kinases of the src-family (e.g., src, yes, lck, lyn, fyn) and non-receptor linked and cytosolic protein tyrosine kinases, such as the Jak family, members of which mediate signal transduction triggered by the cytokine superfamily of receptors (e.g., the Interleukins, Interferons, GM-CSF, and Leptin).

Because of the wide range of known factors capable of stimulating tyrosine kinase activity, the identification of novel human secreted proteins capable of activating

tyrosine kinase signal transduction pathways are of interest. Therefore, the following protocol is designed to identify those novel human secreted proteins capable of activating the tyrosine kinase signal transduction pathways.

Seed target cells (e.g., primary keratinocytes) at a density of approximately
5 25,000 cells per well in a 96 well Loprodyne Silent Screen Plates purchased from
Nalge Nunc (Naperville, IL). The plates are sterilized with two 30 minute rinses with
100% ethanol, rinsed with water and dried overnight. Some plates are coated for 2 hr
with 100 ml of cell culture grade type I collagen (50 mg/ml), gelatin (2%) or polylysine
(50 mg/ml), all of which can be purchased from Sigma Chemicals (St. Louis, MO) or
10 10% Matrigel purchased from Becton Dickinson (Bedford, MA), or calf serum, rinsed
with PBS and stored at 4°C. Cell growth on these plates is assayed by seeding 5,000
cells/well in growth medium and indirect quantitation of cell number through use of
AlamarBlue as described by the manufacturer Alamar Biosciences, Inc. (Sacramento,
CA) after 48 hr. Falcon plate covers #3071 from Becton Dickinson (Bedford, MA) are
15 used to cover the Loprodyne Silent Screen Plates. Falcon Microtest III cell culture
plates can also be used in some proliferation experiments.

To prepare extracts, A431 cells are seeded onto the nylon membranes of
Loprodyne plates (20,000/200ml/well) and cultured overnight in complete medium.
Cells are quiesced by incubation in serum-free basal medium for 24 hr. After 5-20
20 minutes treatment with EGF (60ng/ml) or 50 ul of the supernatant produced in Example
11, the medium was removed and 100 ml of extraction buffer ((20 mM HEPES pH
7.5, 0.15 M NaCl, 1% Triton X-100, 0.1% SDS, 2 mM Na₃VO₄, 2 mM Na₄P₂O₇
and a cocktail of protease inhibitors (# 1836170) obtained from Boehringer Mannheim
(Indianapolis, IN) is added to each well and the plate is shaken on a rotating shaker for
25 5 minutes at 4°C. The plate is then placed in a vacuum transfer manifold and the extract
filtered through the 0.45 mm membrane bottoms of each well using house vacuum.
Extracts are collected in a 96-well catch/assay plate in the bottom of the vacuum
manifold and immediately placed on ice. To obtain extracts clarified by centrifugation,
the content of each well, after detergent solubilization for 5 minutes, is removed and
30 centrifuged for 15 minutes at 4°C at 16,000 x g.

Test the filtered extracts for levels of tyrosine kinase activity. Although many
methods of detecting tyrosine kinase activity are known, one method is described here.

Generally, the tyrosine kinase activity of a supernatant is evaluated by
determining its ability to phosphorylate a tyrosine residue on a specific substrate (a
35 biotinylated peptide). Biotinylated peptides that can be used for this purpose include
PSK1 (corresponding to amino acids 6-20 of the cell division kinase cdc2-p34) and

PSK2 (corresponding to amino acids 1-17 of gastrin). Both peptides are substrates for a range of tyrosine kinases and are available from Boehringer Mannheim.

The tyrosine kinase reaction is set up by adding the following components in order. First, add 10ul of 5uM Biotinylated Peptide, then 10ul ATP/Mg₂⁺ (5mM
5 ATP/50mM MgCl₂), then 10ul of 5x Assay Buffer (40mM imidazole hydrochloride, pH7.3, 40 mM beta-glycerophosphate, 1mM EGTA, 100mM MgCl₂, 5 mM MnCl₂, 0.5 mg/ml BSA), then 5ul of Sodium Vanadate(1mM), and then 5ul of water. Mix the components gently and preincubate the reaction mix at 30°C for 2 min. Initiate the reaction by adding 10ul of the control enzyme or the filtered supernatant.

10 The tyrosine kinase assay reaction is then terminated by adding 10 ul of 120mM EDTA and place the reactions on ice.

Tyrosine kinase activity is determined by transferring 50 ul aliquot of reaction mixture to a microtiter plate (MTP) module and incubating at 37°C for 20 min. This allows the streptavidin coated 96 well plate to associate with the biotinylated peptide.
15 Wash the MTP module with 300ul/well of PBS four times. Next add 75 ul of anti-phosphotyrosine antibody conjugated to horse radish peroxidase(anti-P-Tyr-POD(0.5u/ml)) to each well and incubate at 37°C for one hour. Wash the well as above.

20 Next add 100ul of peroxidase substrate solution (Boehringer Mannheim) and incubate at room temperature for at least 5 mins (up to 30 min). Measure the absorbance of the sample at 405 nm by using ELISA reader. The level of bound peroxidase activity is quantitated using an ELISA reader and reflects the level of tyrosine kinase activity.

25 **Example 20: High-Throughput Screening Assay Identifying Phosphorylation Activity**

As a potential alternative and/or complement to the assay of protein tyrosine kinase activity described in Example 19, an assay which detects activation (phosphorylation) of major intracellular signal transduction intermediates can also be
30 used. For example, as described below one particular assay can detect tyrosine phosphorylation of the Erk-1 and Erk-2 kinases. However, phosphorylation of other molecules, such as Raf, JNK, p38 MAP, Map kinase kinase (MEK), MEK kinase, Src, Muscle specific kinase (MuSK), IRAK, Tec, and Janus, as well as any other phosphoserine, phosphotyrosine, or phosphothreonine molecule, can be detected by
35 substituting these molecules for Erk-1 or Erk-2 in the following assay.

Specifically, assay plates are made by coating the wells of a 96-well ELISA plate with 0.1ml of protein G (1ug/ml) for 2 hr at room temp, (RT). The plates are then rinsed with PBS and blocked with 3% BSA/PBS for 1 hr at RT. The protein G plates are then treated with 2 commercial monoclonal antibodies (100ng/well) against Erk-1 and Erk-2 (1 hr at RT) (Santa Cruz Biotechnology). (To detect other molecules, this step can easily be modified by substituting a monoclonal antibody detecting any of the above described molecules.) After 3-5 rinses with PBS, the plates are stored at 4°C until use.

10 A431 cells are seeded at 20,000/well in a 96-well Loprodyne filterplate and cultured overnight in growth medium. The cells are then starved for 48 hr in basal medium (DMEM) and then treated with EGF (6ng/well) or 50 ul of the supernatants obtained in Example 11 for 5-20 minutes. The cells are then solubilized and extracts filtered directly into the assay plate.

15 After incubation with the extract for 1 hr at RT, the wells are again rinsed. As a positive control, a commercial preparation of MAP kinase (10ng/well) is used in place of A431 extract. Plates are then treated with a commercial polyclonal (rabbit) antibody (1ug/ml) which specifically recognizes the phosphorylated epitope of the Erk-1 and Erk-2 kinases (1 hr at RT). This antibody is biotinylated by standard procedures. The bound polyclonal antibody is then quantitated by successive incubations with
20 Europium-streptavidin and Europium fluorescence enhancing reagent in the Wallac DELFIA instrument (time-resolved fluorescence). An increased fluorescent signal over background indicates a phosphorylation.

Example 21: Method of Determining Alterations in a Gene
25 **Corresponding to a Polynucleotide**

RNA isolated from entire families or individual patients presenting with a phenotype of interest (such as a disease) is be isolated. cDNA is then generated from these RNA samples using protocols known in the art. (See, Sambrook.) The cDNA is then used as a template for PCR, employing primers surrounding regions of interest in
30 SEQ ID NO:X. Suggested PCR conditions consist of 35 cycles at 95°C for 30 seconds; 60-120 seconds at 52-58°C; and 60-120 seconds at 70°C, using buffer solutions described in Sidransky, D., et al., Science 252:706 (1991).

PCR products are then sequenced using primers labeled at their 5' end with T4 polynucleotide kinase, employing SequiTherm Polymerase. (Epicentre Technologies).
35 The intron-exon borders of selected exons is also determined and genomic PCR

products analyzed to confirm the results. PCR products harboring suspected mutations is then cloned and sequenced to validate the results of the direct sequencing.

5 PCR products is cloned into T-tailed vectors as described in Holton, T.A. and Graham, M.W., *Nucleic Acids Research*, 19:1156 (1991) and sequenced with T7 polymerase (United States Biochemical). Affected individuals are identified by mutations not present in unaffected individuals.

10 Genomic rearrangements are also observed as a method of determining alterations in a gene corresponding to a polynucleotide. Genomic clones isolated according to Example 2 are nick-translated with digoxigenin deoxy-uridine 5'-triphosphate (Boehringer Mannheim), and FISH performed as described in Johnson, Cg. et al., *Methods Cell Biol.* 35:73-99 (1991). Hybridization with the labeled probe is carried out using a vast excess of human cot-1 DNA for specific hybridization to the corresponding genomic locus.

15 Chromosomes are counterstained with 4,6-diamino-2-phenylidole and propidium iodide, producing a combination of C- and R-bands. Aligned images for precise mapping are obtained using a triple-band filter set (Chroma Technology, Brattleboro, VT) in combination with a cooled charge-coupled device camera (Photometrics, Tucson, AZ) and variable excitation wavelength filters. (Johnson, Cv. et al., *Genet. Anal. Tech. Appl.*, 8:75 (1991).) Image collection, analysis and
20 chromosomal fractional length measurements are performed using the ISee Graphical Program System. (Inovision Corporation, Durham, NC.) Chromosome alterations of the genomic region hybridized by the probe are identified as insertions, deletions, and translocations. These alterations are used as a diagnostic marker for an associated disease.

25

Example 22: Method of Detecting Abnormal Levels of a Polypeptide in a Biological Sample

30 A polypeptide of the present invention can be detected in a biological sample, and if an increased or decreased level of the polypeptide is detected, this polypeptide is a marker for a particular phenotype. Methods of detection are numerous, and thus, it is understood that one skilled in the art can modify the following assay to fit their particular needs.

35 For example, antibody-sandwich ELISAs are used to detect polypeptides in a sample, preferably a biological sample. Wells of a microtiter plate are coated with specific antibodies, at a final concentration of 0.2 to 10 ug/ml. The antibodies are either monoclonal or polyclonal and are produced by the method described in Example 10.

The wells are blocked so that non-specific binding of the polypeptide to the well is reduced.

The coated wells are then incubated for > 2 hours at RT with a sample containing the polypeptide. Preferably, serial dilutions of the sample should be used to validate results. The plates are then washed three times with deionized or distilled water to remove unbounded polypeptide.

Next, 50 ul of specific antibody-alkaline phosphatase conjugate, at a concentration of 25-400 ng, is added and incubated for 2 hours at room temperature. The plates are again washed three times with deionized or distilled water to remove unbounded conjugate.

Add 75 ul of 4-methylumbelliferyl phosphate (MUP) or p-nitrophenyl phosphate (NPP) substrate solution to each well and incubate 1 hour at room temperature. Measure the reaction by a microtiter plate reader. Prepare a standard curve, using serial dilutions of a control sample, and plot polypeptide concentration on the X-axis (log scale) and fluorescence or absorbance of the Y-axis (linear scale). Interpolate the concentration of the polypeptide in the sample using the standard curve.

Example 23: Formulating a Polypeptide

The secreted polypeptide composition will be formulated and dosed in a fashion consistent with good medical practice, taking into account the clinical condition of the individual patient (especially the side effects of treatment with the secreted polypeptide alone), the site of delivery, the method of administration, the scheduling of administration, and other factors known to practitioners. The "effective amount" for purposes herein is thus determined by such considerations.

As a general proposition, the total pharmaceutically effective amount of secreted polypeptide administered parenterally per dose will be in the range of about 1 $\mu\text{g/kg/day}$ to 10 mg/kg/day of patient body weight, although, as noted above, this will be subject to therapeutic discretion. More preferably, this dose is at least 0.01 mg/kg/day , and most preferably for humans between about 0.01 and 1 mg/kg/day for the hormone. If given continuously, the secreted polypeptide is typically administered at a dose rate of about 1 $\mu\text{g/kg/hour}$ to about 50 $\mu\text{g/kg/hour}$, either by 1-4 injections per day or by continuous subcutaneous infusions, for example, using a mini-pump. An intravenous bag solution may also be employed. The length of treatment needed to observe changes and the interval following treatment for responses to occur appears to vary depending on the desired effect.

Pharmaceutical compositions containing the secreted protein of the invention are administered orally, rectally, parenterally, intracisternally, intravaginally,

intraperitoneally, topically (as by powders, ointments, gels, drops or transdermal patch), buccally, or as an oral or nasal spray. "Pharmaceutically acceptable carrier" refers to a non-toxic solid, semisolid or liquid filler, diluent, encapsulating material or formulation auxiliary of any type. The term "parenteral" as used herein refers to modes of administration which include intravenous, intramuscular, intraperitoneal, intrasternal, subcutaneous and intraarticular injection and infusion.

The secreted polypeptide is also suitably administered by sustained-release systems. Suitable examples of sustained-release compositions include semi-permeable polymer matrices in the form of shaped articles, e.g., films, or microcapsules. Sustained-release matrices include polylactides (U.S. Pat. No. 3,773,919, EP 58,481), copolymers of L-glutamic acid and gamma-ethyl-L-glutamate (Sidman, U. et al., *Biopolymers* 22:547-556 (1983)), poly (2-hydroxyethyl methacrylate) (R. Langer et al., *J. Biomed. Mater. Res.* 15:167-277 (1981), and R. Langer, *Chem. Tech.* 12:98-105 (1982)), ethylene vinyl acetate (R. Langer et al.) or poly-D-(-)-3-hydroxybutyric acid (EP 133,988). Sustained-release compositions also include liposomally entrapped polypeptides. Liposomes containing the secreted polypeptide are prepared by methods known per se: DE 3,218,121; Epstein et al., *Proc. Natl. Acad. Sci. USA* 82:3688-3692 (1985); Hwang et al., *Proc. Natl. Acad. Sci. USA* 77:4030-4034 (1980); EP 52,322; EP 36,676; EP 88,046; EP 143,949; EP 142,641; Japanese Pat. Appl. 83-118008; U.S. Pat. Nos. 4,485,045 and 4,544,545; and EP 102,324. Ordinarily, the liposomes are of the small (about 200-800 Angstroms) unilamellar type in which the lipid content is greater than about 30 mol. percent cholesterol, the selected proportion being adjusted for the optimal secreted polypeptide therapy.

For parenteral administration, in one embodiment, the secreted polypeptide is formulated generally by mixing it at the desired degree of purity, in a unit dosage injectable form (solution, suspension, or emulsion), with a pharmaceutically acceptable carrier, i.e., one that is non-toxic to recipients at the dosages and concentrations employed and is compatible with other ingredients of the formulation. For example, the formulation preferably does not include oxidizing agents and other compounds that are known to be deleterious to polypeptides.

Generally, the formulations are prepared by contacting the polypeptide uniformly and intimately with liquid carriers or finely divided solid carriers or both. Then, if necessary, the product is shaped into the desired formulation. Preferably the carrier is a parenteral carrier, more preferably a solution that is isotonic with the blood of the recipient. Examples of such carrier vehicles include water, saline, Ringer's solution, and dextrose solution. Non-aqueous vehicles such as fixed oils and ethyl oleate are also useful herein, as well as liposomes.

The carrier suitably contains minor amounts of additives such as substances that enhance isotonicity and chemical stability. Such materials are non-toxic to recipients at the dosages and concentrations employed, and include buffers such as phosphate, citrate, succinate, acetic acid, and other organic acids or their salts; antioxidants such as
5 ascorbic acid; low molecular weight (less than about ten residues) polypeptides, e.g., polyarginine or tripeptides; proteins, such as serum albumin, gelatin, or immunoglobulins; hydrophilic polymers such as polyvinylpyrrolidone; amino acids, such as glycine, glutamic acid, aspartic acid, or arginine; monosaccharides, disaccharides, and other carbohydrates including cellulose or its derivatives, glucose,
10 manose, or dextrans; chelating agents such as EDTA; sugar alcohols such as mannitol or sorbitol; counterions such as sodium; and/or nonionic surfactants such as polysorbates, poloxamers, or PEG.

The secreted polypeptide is typically formulated in such vehicles at a concentration of about 0.1 mg/ml to 100 mg/ml, preferably 1-10 mg/ml, at a pH of
15 about 3 to 8. It will be understood that the use of certain of the foregoing excipients, carriers, or stabilizers will result in the formation of polypeptide salts.

Any polypeptide to be used for therapeutic administration can be sterile. Sterility is readily accomplished by filtration through sterile filtration membranes (e.g., 0.2 micron membranes). Therapeutic polypeptide compositions generally are placed
20 into a container having a sterile access port, for example, an intravenous solution bag or vial having a stopper pierceable by a hypodermic injection needle.

Polypeptides ordinarily will be stored in unit or multi-dose containers, for example, sealed ampoules or vials, as an aqueous solution or as a lyophilized formulation for reconstitution. As an example of a lyophilized formulation, 10-ml vials
25 are filled with 5 ml of sterile-filtered 1% (w/v) aqueous polypeptide solution, and the resulting mixture is lyophilized. The infusion solution is prepared by reconstituting the lyophilized polypeptide using bacteriostatic Water-for-Injection.

The invention also provides a pharmaceutical pack or kit comprising one or more containers filled with one or more of the ingredients of the pharmaceutical
30 compositions of the invention. Associated with such container(s) can be a notice in the form prescribed by a governmental agency regulating the manufacture, use or sale of pharmaceuticals or biological products, which notice reflects approval by the agency of manufacture, use or sale for human administration. In addition, the polypeptides of the present invention may be employed in conjunction with other therapeutic compounds.

35

Example 24: Method of Treating Decreased Levels of the Polypeptide

It will be appreciated that conditions caused by a decrease in the standard or normal expression level of a secreted protein in an individual can be treated by administering the polypeptide of the present invention, preferably in the secreted form. Thus, the invention also provides a method of treatment of an individual in need of an
5 increased level of the polypeptide comprising administering to such an individual a pharmaceutical composition comprising an amount of the polypeptide to increase the activity level of the polypeptide in such an individual.

For example, a patient with decreased levels of a polypeptide receives a daily dose 0.1-100 ug/kg of the polypeptide for six consecutive days. Preferably, the
10 polypeptide is in the secreted form. The exact details of the dosing scheme, based on administration and formulation, are provided in Example 23.

Example 25: Method of Treating Increased Levels of the Polypeptide

Antisense technology is used to inhibit production of a polypeptide of the
15 present invention. This technology is one example of a method of decreasing levels of a polypeptide, preferably a secreted form, due to a variety of etiologies, such as cancer.

For example, a patient diagnosed with abnormally increased levels of a polypeptide is administered intravenously antisense polynucleotides at 0.5, 1.0, 1.5, 2.0 and 3.0 mg/kg day for 21 days. This treatment is repeated after a 7-day rest period
20 if the treatment was well tolerated. The formulation of the antisense polynucleotide is provided in Example 23.

Example 26: Method of Treatment Using Gene Therapy

One method of gene therapy transplants fibroblasts, which are capable of
25 expressing a polypeptide, onto a patient. Generally, fibroblasts are obtained from a subject by skin biopsy. The resulting tissue is placed in tissue-culture medium and separated into small pieces. Small chunks of the tissue are placed on a wet surface of a tissue culture flask, approximately ten pieces are placed in each flask. The flask is turned upside down, closed tight and left at room temperature over night. After 24
30 hours at room temperature, the flask is inverted and the chunks of tissue remain fixed to the bottom of the flask and fresh media (e.g., Ham's F12 media, with 10% FBS, penicillin and streptomycin) is added. The flasks are then incubated at 37°C for approximately one week.

At this time, fresh media is added and subsequently changed every several days.
35 After an additional two weeks in culture, a monolayer of fibroblasts emerge. The monolayer is trypsinized and scaled into larger flasks.

pMV-7 (Kirschmeier, P.T. et al., DNA, 7:219-25 (1988)), flanked by the long terminal repeats of the Moloney murine sarcoma virus, is digested with EcoRI and HindIII and subsequently treated with calf intestinal phosphatase. The linear vector is fractionated on agarose gel and purified, using glass beads.

5 The cDNA encoding a polypeptide of the present invention can be amplified using PCR primers which correspond to the 5' and 3' end sequences respectively as set forth in Example 1. Preferably, the 5' primer contains an EcoRI site and the 3' primer includes a HindIII site. Equal quantities of the Moloney murine sarcoma virus linear backbone and the amplified EcoRI and HindIII fragment are added together, in the
10 presence of T4 DNA ligase. The resulting mixture is maintained under conditions appropriate for ligation of the two fragments. The ligation mixture is then used to transform bacteria HB101, which are then plated onto agar containing kanamycin for the purpose of confirming that the vector has the gene of interest properly inserted.

15 The amphotropic pA317 or GP+am12 packaging cells are grown in tissue culture to confluent density in Dulbecco's Modified Eagles Medium (DMEM) with 10% calf serum (CS), penicillin and streptomycin. The MSV vector containing the gene is then added to the media and the packaging cells transduced with the vector. The packaging cells now produce infectious viral particles containing the gene (the packaging cells are now referred to as producer cells).

20 Fresh media is added to the transduced producer cells, and subsequently, the media is harvested from a 10 cm plate of confluent producer cells. The spent media, containing the infectious viral particles, is filtered through a millipore filter to remove detached producer cells and this media is then used to infect fibroblast cells. Media is removed from a sub-confluent plate of fibroblasts and quickly replaced with the media
25 from the producer cells. This media is removed and replaced with fresh media. If the titer of virus is high, then virtually all fibroblasts will be infected and no selection is required. If the titer is very low, then it is necessary to use a retroviral vector that has a selectable marker, such as neo or his. Once the fibroblasts have been efficiently infected, the fibroblasts are analyzed to determine whether protein is produced.

30 The engineered fibroblasts are then transplanted onto the host, either alone or after having been grown to confluence on cytodex 3 microcarrier beads.

It will be clear that the invention may be practiced otherwise than as particularly described in the foregoing description and examples. Numerous modifications and variations of the present invention are possible in light of the above teachings and,
35 therefore, are within the scope of the appended claims.

The entire disclosure of each document cited (including patents, patent applications, journal articles, abstracts, laboratory manuals, books, or other

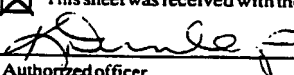
disclosures) in the Background of the Invention, Detailed Description, and Examples is hereby incorporated herein by reference. Further, the hard copy of the sequence listing submitted herewith and the corresponding computer readable form are both incorporated herein by reference in their entireties.

Applicant's or agent's file reference number	PZ021PCT	International application No.	unassigned
---	----------	-------------------------------	------------

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

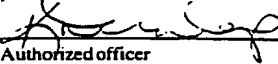
A. The indications made below relate to the microorganism referred to in the description on page <u>172</u> , line <u>N/A</u>	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input type="checkbox"/>	
Name of depositary institution American Type Culture Collection	
Address of depositary institution (including postal code and country) 10801 University Boulevard Manassas, Virginia 20110-2209 United States of America	
Date of deposit 14 November 1997	Accession Number 209463
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications e.g., "Accession Number of Deposit")	

For receiving Office use only <input checked="" type="checkbox"/> This sheet was received with the international application  Authorized officer	For International Bureau use only <input type="checkbox"/> This sheet was received by the International Bureau on: Authorized officer
---	--

Applicant's or agent's file reference number	PZ021PCT	International application No	unassigned
---	----------	------------------------------	------------

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

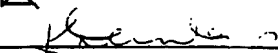
(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>174</u> . line <u>N/A</u>	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input type="checkbox"/>	
Name of depositary institution American Type Culture Collection	
Address of depositary institution (including postal code and country) 10801 University Boulevard Manassas, Virginia 20110-2209 United States of America	
Date of deposit 18 May 1998	Accession Number 209877
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications e.g., "Accession Number of Deposits")	
For receiving Office use only	For International Bureau use only
<input checked="" type="checkbox"/> This sheet was received with the international application  Authorized officer	<input type="checkbox"/> This sheet was received by the International Bureau on: Authorized officer

Applicant's or agent's file reference number	PZ021PCT	International application No	unassigned
---	----------	------------------------------	------------

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

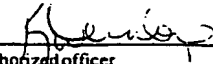
(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>177</u> , line <u>N/A</u>	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input type="checkbox"/>	
Name of depositary institution American Type Culture Collection	
Address of depositary institution (including postal code and country) 10801 University Boulevard Manassas, Virginia 20110-2209 United States of America	
Date of deposit 3 December 1997	Accession Number 209511
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications e.g., "Accession Number of Deposit")	
For receiving Office use only <input checked="" type="checkbox"/> This sheet was received with the international application  Authorized officer	For International Bureau use only <input type="checkbox"/> This sheet was received by the International Bureau on: Authorized officer

Applicant's or agent's file reference number	PZ021PCT	International application No.	unassigned
---	----------	-------------------------------	------------

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

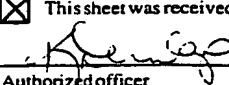
(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>181</u> , line <u>N/A</u>	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input type="checkbox"/>	
Name of depositary institution American Type Culture Collection	
Address of depositary institution (including postal code and country) 10801 University Boulevard Manassas, Virginia 20110-2209 United States of America	
Date of deposit 12 December 1997	Accession Number 209551
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications e.g., "Accession Number of Deposit")	
For receiving Office use only <input checked="" type="checkbox"/> This sheet was received with the international application  Authorized officer	For International Bureau use only <input type="checkbox"/> This sheet was received by the International Bureau on: Authorized officer

Applicant's or agent's file reference number	PZ021PCT	International application No	unassigned
---	----------	------------------------------	------------

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>182</u> , line <u>N/A</u>	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input type="checkbox"/>	
Name of depositary institution American Type Culture Collection	
Address of depositary institution (including postal code and country) 10801 University Boulevard Manassas, Virginia 20110-2209 United States of America	
Date of deposit 7 May 1998	Accession Number 209852
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications e.g., "Accession Number of Deposit")	
For receiving Office use only	For International Bureau use only
<input checked="" type="checkbox"/> This sheet was received with the international application  Authorized officer	<input type="checkbox"/> This sheet was received by the International Bureau on: Authorized officer

What Is Claimed Is:

1. An isolated nucleic acid molecule comprising a polynucleotide having a nucleotide sequence at least 95% identical to a sequence selected from the group consisting of:
 - (a) a polynucleotide fragment of SEQ ID NO:X or a polynucleotide fragment of the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X;
 - (b) a polynucleotide encoding a polypeptide fragment of SEQ ID NO:Y or a polypeptide fragment encoded by the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X;
 - (c) a polynucleotide encoding a polypeptide domain of SEQ ID NO:Y or a polypeptide domain encoded by the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X;
 - (d) a polynucleotide encoding a polypeptide epitope of SEQ ID NO:Y or a polypeptide epitope encoded by the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X;
 - (e) a polynucleotide encoding a polypeptide of SEQ ID NO:Y or the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X, having biological activity;
 - (f) a polynucleotide which is a variant of SEQ ID NO:X;
 - (g) a polynucleotide which is an allelic variant of SEQ ID NO:X;
 - (h) a polynucleotide which encodes a species homologue of the SEQ ID NO:Y;
 - (i) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(h), wherein said polynucleotide does not hybridize under stringent conditions to a nucleic acid molecule having a nucleotide sequence of only A residues or of only T residues.
2. The isolated nucleic acid molecule of claim 1, wherein the polynucleotide fragment comprises a nucleotide sequence encoding a secreted protein.
3. The isolated nucleic acid molecule of claim 1, wherein the polynucleotide fragment comprises a nucleotide sequence encoding the sequence identified as SEQ ID NO:Y or the polypeptide encoded by the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X.

4. The isolated nucleic acid molecule of claim 1, wherein the polynucleotide fragment comprises the entire nucleotide sequence of SEQ ID NO:X or the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X.
5. The isolated nucleic acid molecule of claim 2, wherein the nucleotide sequence comprises sequential nucleotide deletions from either the C-terminus or the N-terminus.
6. The isolated nucleic acid molecule of claim 3, wherein the nucleotide sequence comprises sequential nucleotide deletions from either the C-terminus or the N-terminus.
7. A recombinant vector comprising the isolated nucleic acid molecule of claim 1.
8. A method of making a recombinant host cell comprising the isolated nucleic acid molecule of claim 1.
9. A recombinant host cell produced by the method of claim 8.
10. The recombinant host cell of claim 9 comprising vector sequences.
11. An isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence selected from the group consisting of:
 - (a) a polypeptide fragment of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z;
 - (b) a polypeptide fragment of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z, having biological activity;
 - (c) a polypeptide domain of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z;
 - (d) a polypeptide epitope of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z;
 - (e) a secreted form of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z;
 - (f) a full length protein of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z;

- (g) a variant of SEQ ID NO:Y;
 - (h) an allelic variant of SEQ ID NO:Y; or
 - (i) a species homologue of the SEQ ID NO:Y.
12. The isolated polypeptide of claim 11, wherein the secreted form or the full length protein comprises sequential amino acid deletions from either the C-terminus or the N-terminus.
13. An isolated antibody that binds specifically to the isolated polypeptide of claim 11.
14. A recombinant host cell that expresses the isolated polypeptide of claim 11.
15. A method of making an isolated polypeptide comprising:
- (a) culturing the recombinant host cell of claim 14 under conditions such that said polypeptide is expressed; and
 - (b) recovering said polypeptide.
16. The polypeptide produced by claim 15.
17. A method for preventing, treating, or ameliorating a medical condition, comprising administering to a mammalian subject a therapeutically effective amount of the polypeptide of claim 11 or the polynucleotide of claim 1.
18. A method of diagnosing a pathological condition or a susceptibility to a pathological condition in a subject comprising:
- (a) determining the presence or absence of a mutation in the polynucleotide of claim 1; and
 - (b) diagnosing a pathological condition or a susceptibility to a pathological condition based on the presence or absence of said mutation.
19. A method of diagnosing a pathological condition or a susceptibility to a pathological condition in a subject comprising:
- (a) determining the presence or amount of expression of the polypeptide of claim 11 in a biological sample; and
 - (b) diagnosing a pathological condition or a susceptibility to a pathological condition based on the presence or amount of expression of the polypeptide.

20. A method for identifying a binding partner to the polypeptide of claim 11 comprising:

- (a) contacting the polypeptide of claim 11 with a binding partner; and
- (b) determining whether the binding partner effects an activity of the polypeptide.

21. The gene corresponding to the cDNA sequence of SEQ ID NO:Y.

22. A method of identifying an activity in a biological assay, wherein the method comprises:

- (a) expressing SEQ ID NO:X in a cell;
- (b) isolating the supernatant;
- (c) detecting an activity in a biological assay; and
- (d) identifying the protein in the supernatant having the activity.

23. The product produced by the method of claim 20.

<110> Human Genome Sciences, Inc. et al
<120> 110 Human Secreted Proteins
<130> PZ021PCT
<140> PCT/US98/27059
<141> 1998-12-17

<150> 60/068,008
<151> 1998-12-18

<150> 60/068,054
<151> 1998-12-18

<150> 60/068,064
<151> 1998-12-18

<150> 60/068,053
<151> 1998-12-18

<150> 60/068,006
<151> 1998-12-18

<150> 60/068,057
<151> 1998-12-18

<150> 60/068,007
<151> 1998-12-18

<150> 60/070,923
<151> 1998-12-18

<150> 60/068,367
<151> 1998-12-19

<150> 60/068,369
<151> 1998-12-19

<150> 60/068,169
<151> 1998-12-19

<150> 60/068,365
<151> 1998-12-19

<150> 60/068,368
<151> 1998-12-19

<160> 628

<170> PatentIn Ver. 2.0

<210> 1
<211> 733
<212> DNA

<213> Homo sapiens

<400> 1

gggatccgga	gcccaaattct	tctgacaaaa	ctcacacatg	cccaccgtgc	ccagcacctg	60
aattcgaggg	tgcaccgtca	gtcttcctct	tcccccaaa	acccaaggac	accctcatga	120
tctcccgga	tcttgaggtc	acatgcgtgg	tggtggacgt	aagccacgaa	gaccctgagg	180
tcaagttcaa	ctggtacgtg	gacggcgtgg	aggtgcataa	tgccaagaca	aagccgcggg	240
aggagcagta	caacagcacg	taccgtgtgg	tcagcgtcct	caccgtcctg	caccaggact	300
ggctgaatgg	caaggagtac	aagtgcaagg	tctccaacaa	agccctccca	accccatcgc	360
agaaaaccat	ctccaaagcc	aaagggcagc	cccgagaacc	acaggtgtac	accctgcccc	420
catccccgga	tgagctgacc	aagaaccagg	tcagcctgac	ctgcctggtc	aaaggcttct	480
atccaagcga	catcgccgtg	gagtgggaga	gcaatgggca	gccggagaac	aactacaaga	540
ccacgcctcc	cgtgctggac	tccgacggct	ccttcttctt	ctacagcaag	ctcaccgtgg	600
acaagagcag	gtggcagcag	gggaacgtct	tctcatgctc	cgtgatgcat	gaggctctgc	660
acaaccacta	cacgcagaag	agcctctccc	tgtctccggg	taaatgagtg	cgacggccgc	720
gactctagag	gat					733

<210> 2

<211> 5

<212> PRT

<213> Homo sapiens

<220>

<221> Site

<222> (3)

<223> Xaa equals any of the twenty naturally occurring L-amino acids

<400> 2

Trp	Ser	Xaa	Trp	Ser
1				5

<210> 3

<211> 86

<212> DNA

<213> Homo sapiens

<400> 3

gcgcctcgag	atttccccga	aatctagatt	tccccgaaat	gatttccccg	aaatgatttc	60
cccgaatat	ctgccatctc	aattag				86

<210> 4

<211> 27

<212> DNA

<213> Homo sapiens

<400> 4

gcggcaagct	ttttgcaaag	cctaggc	27
------------	------------	---------	----

<210> 5

<211> 271

<212> DNA

<213> Homo sapiens

<400> 5

ctcgagattt	ccccgaaatc	tagatttccc	cgaaatgatt	tccccgaaat	gatttccccg	60
aaatatctgc	catctcaatt	agtcagcaac	catagtcccc	cccctaactc	cgcccatccc	120
gcccctaact	ccgcccagtt	ccgcccattc	tccgccccat	ggctgactaa	ttttttttat	180
ttatgcagag	gccgaggccg	cctcggcctc	tgagctattc	cagaagtagt	gaggaggcct	240
ttttggaggc	ctaggctttt	gcaaaaagct	t			271

<210> 6
 <211> 32
 <212> DNA
 <213> Homo sapiens

<400> 6
 gcgctcgagg gatgacagcg atagaacccc gg 32

<210> 7
 <211> 31
 <212> DNA
 <213> Homo sapiens

<400> 7
 gcgaagcttc gcgactcccc ggatccgcct c 31

<210> 8
 <211> 12
 <212> DNA
 <213> Homo sapiens

<400> 8
 ggggactttc cc 12

<210> 9
 <211> 73
 <212> DNA
 <213> Homo sapiens

<400> 9
 gcggcctcga ggggactttc ccggggactt tccggggact ttccgggact ttccatcctg 60
 ccatctcaat tag 73

<210> 10
 <211> 256
 <212> DNA
 <213> Homo sapiens

<400> 10
 ctcgagggga ctttcccgga gactttccgg ggactttccg ggactttcca tctgccatct 60
 caattagta gcaaccatag tcccgcctt aactccgcc atcccgcctt taactccgcc 120
 cagttccgcc cattctccgc cccatggctg actaattttt tttatttatg cagaggccga 180
 ggccgcctcg gcctctgagc tattccagaa gtagtgagga ggcttttttg gaggcctagg 240
 cttttgcaaa aagctt 256

<210> 11
 <211> 1271
 <212> DNA
 <213> Homo sapiens

<400> 11
 ggggctgggc cctgctcagg tggctctctc cttgcagggg cgggcatgc tctgcaggct 60
 gtgctggctg gtctcgtaga gcttggtgtg gctgttgctc ggctgcctgc tcttcctgag 120
 gaaggcggcc aagcccgag agacccacg gccaccagc ctttctgggg ctccccaac 180
 acccggtcac agccggtgtc caccacaacca cacagtgtct agcgctctc tgtccctgcc 240
 tagccgtcac cgtctcttct tgacctatcg tcaactgccg aatttctcta tcttgcaggga 300
 gccttcaggc tgttccaagg ataccttctt gctcctggcc atcaagtcac agcctgggtca 360
 cgtggagcga cgtgcggcta tccgcagcac gtggggcagg tgggggatg ggctaggggc 420
 ggactgaag ctggtgttcc tcctaggggt ggcaggatcc gctccccag ccagctgct 480
 ggcttatgag agtagggagt ttgatgacat cctccagtgg gacttcaact aggacttctt 540
 caacctgacg ctcaaggagc tgcacctgca gcgctgggtg gtggctgcct gccccaggc 600
 ccatttcatg ctaaaggagg atgacgatgt cttgtccac gtcccaacg tgttagagtt 660
 cctggatggc tgggaccag cccaggacct cctggtggga gatgtcatcc gccaaagcct 720
 gccaacagg aacactaagg tcaaatactt catccaccc tcaatgtaca gggccacca 780
 ctaccaccc tatgctggtg ggggaggata tgtcatgtcc agagccacag tggggcgcct 840
 ccaggctatc atggaagatg ctgaactctt cccattgat gatgtctttg tgggtatgtg 900
 cctgaggagg ctggggctga gccctatgca ccattgctgg ttcaagacat ttggaatccg 960
 gcggccctg gacccttag acccctgcct gtatagggg ctctgctgg ttcaccgct 1020
 cagccccctc gagatgtgga ccattgtggc actggtgaca gatgaggggc tcaagtgtgc 1080
 agctggcccc ataccacagc gctgaagggt ggggtgggca acagcctgag agtggactca 1140
 gtgttgattc tctatcgtga tgcgaattg atgcctgctg ctctacagaa aatgccaaact 1200
 tggtttttta actcctctca cctgttagc tctgattaaa aacactgcaa ccaaaaaaa 1260
 aaaaaaaaa a 1271

<210> 12
 <211> 1451
 <212> DNA
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (937)
 <223> n equals a,t,g, or c

<220>
 <221> SITE
 <222> (965)
 <223> n equals a,t,g, or c

<220>
 <221> SITE
 <222> (975)
 <223> n equals a,t,g, or c

<400> 12
 gccagttcac tcctcggctg gagacactag gtgtcctgcc cctaattctca ccagagccac 60
 ctcataaacc ctgcaagctc tgcgaagggc agctgggcac agctgaaagc ccagccacca 120
 gccatgtcc tgggtgggac tgggcaggag gggccacctc ctactgctga tcaaccgag 180
 agccctggct gggatccgtc ttcttcacc aacgggagct cgggccccag ggccctgccc 240
 acctctgtgc accccacact gcagcaggga gcacctgca ggaggaactg ggcaccctgc 300
 aggggtctgg tggagacgag gatgctacgg aggcagctgc cccatgggac cagtaaggag 360

gatcttgggt	gggcttccct	gcagagagga	agccctcagg	agacaccaca	gtaagccatg	420
ctggagacct	ggaggccagg	cccgtcamct	ggggagctgg	ccactaacag	cgggcagaga	480
gcctcccagg	acagccagca	cagcccccca	cacgtccgag	cccacctcct	catttcccgc	540
cttcccgcgt	tcccagcat	gggaggacct	gccggacgca	gcgcaccaty	tytctaaca	600
gagaccaagt	ctgagcttca	acgcttgccg	agacgacagg	cacgtgcaag	cytytctyt	660
ccagcaggag	agcccgagg	aggacacagc	gaytctttca	actgcgtccc	gacaaacggt	720
cagccccctc	gctcctgcag	tttatccaag	ctcaggagga	gctttctaaa	gagaacacag	780
ggagacagct	ggctgccaga	gaagcagtc	tggctctgga	aggctccacc	cagctaacag	840
ggcctgtgac	acaggtggca	gccagcaaga	ccactgcag	cggcattggc	ctcacagcct	900
cccctgtccc	tgtcctggga	gcagccccgg	caaagantcc	taccagaac	aytccaggtc	960
agganggcag	ggccttgarg	aargtgarga	catcctggar	arctgtggcc	accaagggtc	1020
tgcacggcct	ggaggtctcc	acacatctgg	gcaaacggaa	gctttctggg	aggagctggc	1080
tcccaggccc	tgtctctccac	gccaccccat	cacagtgcga	cacacagaca	ggctcccaga	1140
ttgtccaccc	tccacaggga	gaagtcaggg	agggtggcag	gggacgggg	cagccaccgg	1200
ctcagcctgt	gcacgcccac	ccctcccagc	agcacccttc	tccggcccac	ctggctggcc	1260
tgagtctgtg	gactggcact	gcctgataga	actttcagta	ccttcagtgc	ccaaagggcg	1320
cgacgactag	cccttaaaag	aggctggagc	ccctgaggaa	gcgggtcttt	aggcaaaacc	1380
accgcacggc	agagcagcca	caggacagat	caatgtcttc	tgcgggaaaa	aaaaaaaaaa	1440
aaaaactcgt	a					1451

<210> 13

<211> 2317

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (1419)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (2165)

<223> n equals a,t,g, or c

<400> 13

cctgcagcta	ccgtccgcaa	ttcccggctg	accacgcgt	ccgggcgact	tgcctgtct	60
tcaacatgaa	gatagccaca	gtgtcagtgc	ttctgccctt	ggctctttgc	ctcatacaag	120
atgctgccag	taagaatgaa	gatcaggaaa	tgtgccatga	atttcaggca	tttatgaaaa	180
akggaaaact	gttctgtccc	caggataaga	aattttttca	aagtcttgat	ggaataatgt	240
tcatcaataa	atgtgccacg	tgcaaaatga	tactggaaaa	agaagcaaaa	tcacagaaga	300
gggcccaggca	tttagcaaga	gctcccaagg	ctactgcccc	aacagagctg	aattgtgatg	360
attttaaaaa	aggagaaaga	gatggggatt	ttatctgtcc	tgattattat	gaagctgttt	420
gtggcacaga	tgggaaaaca	tatgacaaca	gatgtgcact	gtgtgctgag	aatgcgaaaa	480
ccgggtccca	aattgggtga	aaaagtgaag	gggaatgtaa	gagcagtaat	ccagagcagg	540
atgtatgcag	tgtcttttcg	ccctttgtta	gagatggaa	acttggatgc	acaagggaaa	600
atgatcctgt	tcttggctct	gatgggaaga	cgcatggcaa	taagtgtgca	atgtgtgcts	660
agctgtyyyw	aaaagaagct	gaaaatgcca	agcgagaggg	tgaaactaga	attcgacgaa	720
atgctgaaaa	ggatttttgc	aaggaaatwtg	aaaaacaagt	gagaaatgga	aggctttttt	780
gtacacggga	gagtgatcca	gtccgtggcc	ctgacggcag	gatgcattgg	aacaaatgtg	840
ccctgtgtgc	tgaaattttc	aagcagcggt	tttcagagga	aaacagtaaa	acagatcaaa	900
atltgggaaa	agctgaagaa	aaaactaaag	ttaaaagaga	aattgtgaaa	ctctgcagtc	960
aatatcaaaa	tcaggcaaa	aatggaatac	ttttctgtac	cagagaaaaat	gaccctattc	1020
gtgggtccaga	tgggaaaatg	catggcaact	tgtgttccat	gtgtcaagcc	tacttccaag	1080
cagaaaatga	agaaaagaaa	aaggctgaag	cacgagctag	aaacaaaaga	gaatctggaa	1140
aagcaacctc	atatgcagag	ctttgcagtg	aatatcgaaa	gcttgtgagg	aacggaaaac	1200

SUBSTITUTE SHEET (RULE 26)

ttgcttgcac	cagagagaac	aatcctatcc	agggcccaga	tgggaaagt	catggcaaca	1260
cctgctccat	gtgtgaggtc	ttcttccaag	cagaagaaga	agaaaagaaa	aagaaggaag	1320
gtwmmtcaag	aaacaaaaga	caatctaaga	gtacagcttc	ctttgwkag	ttgtgtagtg	1380
aatmccgcaa	atccaggaaa	aacggacggc	ttttttgcnc	cagagagaat	gaccccatcc	1440
agggcccaga	tggaaaaatg	catggcaaca	cctgctccat	gtgtgaggcc	ttctttcaac	1500
aagaagaaag	agcaagagca	aaggctaaaa	gagaagctgc	aaaggaaaatc	tgcaagtgaat	1560
ttcgggacca	agtgaggaat	ggaacactta	tatgcaccag	ggagcataat	cctgtccgtg	1620
gcccagatgg	caaaatgcat	ggaacaagt	gtgccatgtg	tgccagtgtg	ttcaaaacttg	1680
aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aagggcggcc	gctctagagg	atccctcgag	1740
gggcccagc	ttacgcgtgc	atgcgacgtc	atagctctct	ccctatagtg	agtcgtatta	1800
taagctaggc	actggccgtc	gttttacaac	gtcgtgactg	ggagatctgc	tagcttgga	1860
tctttgtgaa	ggaaccttac	ttctgtgtgtg	tgacataatt	ggacaaacta	cctacagaga	1920
tttaagctc	taaggtaaat	ataaaatttt	taagtgtata	atgtgttaaa	ctagctgcac	1980
atgcttgctg	cttgagagtt	ttgcttactg	agtatgattt	atgaaaatat	tatacacagg	2040
agctagtgat	tctaattggt	tgtgtatttt	agattcacag	tcccaaggct	catttcaggc	2100
ccctcagtc	tcacagtctg	ttcatgatca	taatcagcca	taccacattt	gtagaggktt	2160
tactngcytt	aaaaaacctc	ccacacctcc	ccctgaacck	gaaacataaa	atggaakgca	2220
atggtgggtg	taactgggtt	atgcagcctt	aataaatggg	ttaccaaatt	aaaggccaat	2280
agggcatcca	ccmaaaattt	tccacccaaa	ttaaaag			2317

<210> 14

<211> 1472

<212> DNA

<213> Homo sapiens

<400> 14

ggccacgcgt	ccgcggtacg	gtggtgcggc	tgccgacgca	cagaccagct	gcctacatcc	60
ttgtctccac	cgtgctaacc	ctcatggtgc	cctggcacag	cctggacccc	gactcagcgc	120
ttgcagatgc	cttctaccag	cggggctaca	ggtgggctgg	cttcacgtg	gcagctggct	180
ccatctgcgc	catgaacacc	gtcctgtctca	gcctcctctt	ctccctgcc	cgcattgtct	240
atgccatggc	cgccgatggg	ctcttcttcc	aggtgtttgc	ccatgtgcac	ccccggacac	300
aggtgcctgt	ggcgggcacc	ctggcgctcg	ggcctctcac	ggccttctcg	gcactgctgc	360
tggacctgga	gtcgctgggt	cagttcctgt	cccttggcac	actcctggcc	tacacattcg	420
tggccaccag	tatcattgtg	ctgcgcttcc	agaagtcttc	cccgccacgc	tcccaggcc	480
cagccagccc	tggccccctg	accaagcagc	agagctcctt	ctcagaccac	ctacagctgg	540
tgggcactgt	acacgcctcc	gtccctgagc	caggggagct	gaagccagcc	ctgaggccct	600
acctgggctt	cttygatggg	tacagccctg	gagcagtggt	gacttgggcg	cttggcggtta	660
tgttggcctc	agccatcacc	ataggctgcg	tgttgtctt	tgggaactcg	accctgcacc	720
tcccacactg	gggttacatc	ctgctgtctc	tgctcaccag	tgtcatgttt	ctgctcagcc	780
tccttgtcct	gggggctcac	cagcaacagt	atcgggaaga	cttattttcag	atccccatgg	840
ttcccctgat	tccagccctg	agcatcgctc	tcaacatctg	cctcatgctg	aaacttagct	900
atctgacctg	ggtgcgcttc	tccatctggc	tgctgatggg	acttgacgtg	tatttcggct	960
atggcatccg	gcatagcaag	gagaaccagc	gggagctgcc	agggctgaac	tccacacact	1020
acgtgggtatt	ccccaggggc	agcctggagg	agacagtgc	ggctatgcag	ccccccagcc	1080
aggcaccagc	acaggaccct	ggccatatgg	agtagctgat	cagccccacac	ttgccccgcc	1140
ctcccacacc	tgcttgggag	gccagagagg	ccagacaagc	cgagagcccc	ttctgtttgtg	1200
ggcagcctgg	gtttgcaggc	ctgcacaggc	tggggagtcc	tcaggacctt	aggaccttca	1260
tccaggggct	gggcttcggg	tcttcaggag	tgggccttgg	ctggtgctgg	tgccatggac	1320
tctgcccaga	gccttcttgt	ttatgatcag	ctccagctac	ctgggcagtt	gtggtggggt	1380
ggatgggaag	gcccacagcc	caagggatcc	ataataataa	ttgcttggcc	agccatgtgg	1440
cctgctggcg	taaaaaaaaa	aaaaaaaaaa	aa			1472

<210> 15

<211> 1016

<212> DNA

<213> Homo sapiens

<400> 15

ggcagcagct	tccgcggcat	gatttccacc	cagcccggct	ccaccccact	cgcttccctt	60
aagatcctgg	ctctggagtc	ggcagatggg	catggcggct	gcagtgcctg	caatgacatt	120
ggccctctacg	gtgagcggga	cgaccagcaa	gtgttcatcc	agaagggtgt	gcccagtgcc	180
agccagctct	tcgtgcgtct	ctcatctact	gggcagcggg	tgtgctccgt	gcgctccgtg	240
gacggctcac	ccacgacagc	cttcacagtg	ctggagtgcg	agggctcccc	ggcggtcctg	300
ctctcggccc	cggcgctacc	tgctcactgg	ccaggccaac	ggcagcttgg	ccatgtggga	360
cctaaccacc	gccatggacg	gcctcggcca	ggccctgca	ggtggcctga	cggagcaaga	420
gctgatggaa	cagctggaac	actgtgagct	ggcccgcgg	gctcctttca	gctccctcat	480
ggggctgtct	ccccagcccc	tcaccccgca	tctccctcac	cagcctccac	tcagcctcca	540
gcaacacctc	cttgtctggc	caccgtggga	gcccaagccc	ccgcaggct	gaggcccggc	600
gccgtgggtg	gggcagcttt	gtggaacgct	gccaggaact	ggtgcggagt	gggcccagacc	660
tccgacggcc	acccacacca	gccccgtggc	cctccagcgg	tctcggcact	cccctcacac	720
ctcccaagat	gaagtcfaat	gaaacttcct	ttttgaacaa	cgcagctgcc	atgatgcctt	780
gggatgccct	ggtcctgggg	gactcaggtg	cctccctgat	tcctgtggga	accccggtt	840
caggccaggg	cctccttgga	ataaatgggt	attgttacta	ggtccccacc	ttcctctctt	900
tctggaagcc	aaagtcaccc	tccccaataa	agtcctcact	gccaaaaaaa	aaaaaaaaaa	960
aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaa	1016

<210> 16

<211> 1239

<212> DNA

<213> Homo sapiens

<400> 16

cccacgcgtc	cgcccacgcg	tccgcccacg	cgctccggctg	cgggcgcatg	gcggcggggc	60
tggcgcggtc	cctgttgctc	ctcgggtctc	cgcccgcccg	gcccgcgcg	gcaggtgcag	120
cgaagatgaa	ggtgggtggg	gagcccaacg	cgtttgggg	gaacaacccg	ttcttgctc	180
aggccagtcg	cctccaggcc	aagagggtac	cttcacccgt	gtctggaccc	gtgcatctct	240
tccgactctc	gggcaagtgc	ttcagcctgg	tggagtccac	gtacaagtat	gagttctgcc	300
cgttccacaa	cgtgaccag	cacgagcaga	ccttcgcgtg	gaacgcctac	agtgggatcc	360
tcggcatctg	gcacgagtgg	gagatcgcca	acaacacctt	cacgggcatg	tggatgaggg	420
acggtgacgc	ctgccgttcc	cggagccggc	agagcaaggt	ggagctggcg	tgtggaaaaa	480
gcaaccggct	ggcccatgtg	tccgagccga	gcacctgcgt	ctacgcgctg	acgttcgaga	540
ccccctcgt	ctgccacccc	cacgccttgc	tagtgtaccc	aaccttgcca	gaggccctgc	600
agcggcagtg	ggaccaggta	gagcaggacc	tggccgatga	gctgatcacc	ccccagggcc	660
atgagaagtt	gctgaggaca	ctttttgagg	atgctggcta	cttaaagacc	ccagaagaaa	720
atgaaccac	ccagctggag	ggaggctctg	acagcttggg	gtttgagacc	ctggaaaact	780
gcaggaaggc	tcataaagaa	ctctcaaagg	agatcaaaa	gctgaaaggt	ttgtcaccc	840
agcacggcat	cccctacacg	aggcccacag	aaacttccaa	cttggagcac	ttgggccacg	900
agacgcccag	agccaagtct	ccagagcagc	tgcgggtgga	cccaggactg	cgtgggagtt	960
tgtgaccttg	tgggtggaga	gcagaggtgg	acgcggccga	gagccctaca	gagaagctgg	1020
ctggtaggac	ccgcagggac	cagctgacca	ggcttgtgct	cagagaagca	gacaaaacaa	1080
agattcsagg	ttttaattaa	ttccatact	gataaaaata	actccatgaa	ttctgtaaac	1140
cattgcataa	atgctatagt	gtaaaaaat	ttaaacaggt	gttaacttta	aacagttcgc	1200
tacaagtaaa	tgattataaa	tactaaaaaa	aaaaaaaaaa			1239

<210> 17

<211> 1405

<212> DNA

<213> Homo sapiens

<220>

<221> SITE
 <222> (1403)
 <223> n equals a,t,g, or c

<220>
 <221> SITE
 <222> (1404)
 <223> n equals a,t,g, or c

<220>
 <221> SITE
 <222> (1405)
 <223> n equals a,t,g, or c

<400> 17
 gaattcggca cgaggcagct ttggacatgt cgggcctcat agggagcatg ggggtgtgaa 60
 gtgggtggccg tgggctcaaa agctggctgc ttagtttacc agctgtgtga tctcaagcag 120
 atcaccttct tcttcagagc ctcagtttgc ctgtcgggtca tgccctgcct ggaggccgtg 180
 gccttgatcc tgcttatcct gctgggtcca gaccaccccc ggggagctgc cgagacacag 240
 ggggaggggg ccgtgggagg cttcaggagc agctgggtgtg aggacgtcag atacctgggg 300
 aaaaactgga gtttcgtgtg gtcgayyctc rgagtgaccg ccatggcctt tgtgactgga 360
 gccctggggt tctgggcccc caagtctctg ctcgaggcac gcgtgggtca cgggctgcag 420
 cctccctgct tccaggagcc gtgcagcaac cccgacagcc tgatttttgg ggcaactgacc 480
 atcatgaccg gcgtcattgg ggtcatcttg ggggcagaag ctgcgaggag gtacaagaaa 540
 gtcattccag gagctgagcc cctcatctgc gcctccagcc tgcttgccac agccccctgc 600
 ctctacctgg ctctcgctct ggccccgacc accctgctgg cctcctatgt gttcctgggc 660
 cttggggarc tgcttctgtc ctgcaactgg gcagtgggtg ccgacatcct gctgtctgtg 720
 gtgggtgccc gatgccgggg gacggcagaa gcacttcaga tcacgggtggk ycacatcctg 780
 ggaracctgg cagccctatc tcacaggact tatctctagt gtcctgcggg ccargcgccc 840
 tgactcctat ctgcagcgct tccgcagcct gsarcararc tycctgtgct gcgcctttgt 900
 catcgccctg gggggcggtc gcttctctgt gactgcgctg tacctggaga gagacgagac 960
 ccgggcctgg cagmctgtca cagggacccc agacagcaat gatgtggaca gcaatgacct 1020
 ggagagacaa ggctgcttt cgggcgytgg cgcctctaca gaggagccct gaggtccttg 1080
 cctgcactcg tctgctctgc aagcctcccg ttggtcccca cagcagcagt gcctcggttc 1140
 ctctttggct gtcctcgggg actccggctg aggcacatct gccacttttg aattcccggc 1200
 tggagagctg gcaggaccct gtggctgggc tgggaatgga gctgtcagca ctctgcgtgg 1260
 gaggcctggg cctgtgcctg catcccgctc aaggctgccc cagcctgggg tcccagcct 1320
 ggctgctgct gggccctgra taaagagagg ccagtacaaa gcccatggat tttgggcctg 1380
 taaaaaaaaa aaaaaaaaaa aannn 1405

<210> 18
 <211> 1534
 <212> DNA
 <213> Homo sapiens

<400> 18
 ggcacgagtc aagcaatctg tccacctcag cctcccaaag tgctggaatt acaagcataa 60
 ccactgcacc tggtaggcat caaattttga atagagaagt taaccatgat gaatcaacac 120
 ttgcttgaat cgtttggttc tccctcctcc ttgttcattg tctttattct gctcatctgg 180
 atgttgcaaa gatgtaaaga ttttttccct tggtgctata gagtagtgct aactccatca 240
 ttctggcaga agcaccaaca cccagatccc aaaattaagc atcatttgaa gctatactca 300
 ctgaaataca gttcttctgg gcagaacaac ttcagaaagg acaaacattg gctttctggc 360
 cacacggaag aggcaaattt aataaaggaa gaatggaagt aatgttcaat atcaccagga 420
 tagcctacct gaaatttctt agtacaaatg agaagcaatg tggcccatgt ggctgtata 480
 tgagcacgtg gaaagagtga tcttaagttc tccactgata ctaatctaat ctgtgttatt 540
 cgcgtgtacc cttaattatc catttaaaaa gctaattatt gctgttagtg gtgggtgtgg 600

ttgcatgtat	atacgtgttt	caaaaccatc	tgacacagcc	tattcaatgc	tttctctag	660
ttcatacctc	tgtactgggt	tcctgtgtct	gccacagaaa	ttgccaccaa	ggtagcttaa	720
agtaatggaa	atctgctctc	tcacttttct	ggaggctaca	aatctgcaat	caagggtgca	780
acaggccatg	ctccctctga	aggctctgcg	gaagaatcct	ttcttgcttc	ttcctagttt	840
tgatgggtgc	tgccaatcct	tggcattccc	tggttgtgg	ctgcaacact	ccaatctctg	900
cctcaatcat	cacatgacct	tccttggtga	tctcctttgt	gtctctgtgt	tcaaataatt	960
cttccctttc	tcctgtacat	ataccagtca	ttggatttag	ggttctcttg	atccagtatg	1020
atcttatctc	aactggatta	tatcttcaaa	gaccttattt	gaacgcctta	ttcccataac	1080
ggtcacattt	ccaagtacaa	gggttaggac	ttaaaacata	ccttcttggg	ggccacaatt	1140
taacctatta	tatcccacta	cacagtatat	gccatgacag	aacctctttc	atgagaccaa	1200
catgaaatgt	agagtgatat	gattctttat	gtgatcaatg	agtatttgca	gacaatgctg	1260
atgattcccc	tgagaatatc	atggctccct	cccttactta	tttagatata	tggtcaggct	1320
gggtgctgtg	gctcatgcct	ataatcccag	cactttggga	ggggaggaag	gtgggttcaact	1380
tgagcccagg	agttcaagac	cagcttgggc	aacatggtga	gaccttggtt	ttaccaaaaa	1440
aaagaaagaa	aaatacaaaa	aattagctgg	gtgtggtggc	acatgcctgt	agttcctgct	1500
acttgggtgg	ctgggggtgg	agaatcatct	gagc			1534

<210> 19

<211> 1233

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (491)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (493)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (497)

<223> n equals a,t,g, or c

<400> 19

gcagcgtgag	ccaccgtgcc	tggtctggtt	ggggattttc	atattgcatg	tgacagaaaa	60
cctatctgta	actagctaaa	tagccctttc	ccctactgcc	ccccaaaagg	cggggtttat	120
tgtatcatct	gattcagaaa	tctacgtcgg	gcttggtagt	ttggtttcgg	gaacatctgg	180
attccaggtc	tcaaatgaca	tcacgtctat	tcatttttct	ctttctctgg	ttctgcctc	240
caccacgtat	cagctttggt	ctgtgttggc	ctcagcccca	ttctcaagtt	cacattcagc	300
atgaaaaggc	tgatcacctc	ttccagtcac	tcaagcaaaa	agccccaggt	ttgctgcaat	360
gggctagaat	agtttgactg	taatcacatg	accacctcta	agccaatctc	tgtgggccag	420
gcaccccagt	ggttggttta	rgcctggatc	ttgtgacaca	ctcctgaaca	catgactcga	480
gggtggggca	nangcanatc	ctcacacaga	actggtgcgt	gattccaggt	ggctgatcaa	540
tcacctgtgt	ccactgtggt	tgtatcaarg	gcgtgtgggt	tgtgctgtgc	tgggcgtaca	600
ctagtgtgtg	cattggccag	ctcggkktgc	cacgttattg	gctttgggat	gctggttccc	660
aggagtggcg	gtgggaagat	tccacctgct	ttctatggta	ggaaggcagg	tcctgggggtg	720
agggtgagcc	ccgaggggga	ccagtggcca	ctgtggcttg	acccgcaggc	cttgacctga	780
gcattgcagg	catagtttcc	tgcccctgta	actgccagat	cgagaccag	tgagacagtg	840
gtttccgctg	ccagagactg	aaatagtgtg	ttccctgagg	acccctgaag	aagcttggct	900
ttgtacagag	ccagacgtcc	aggtcacaa	tctggctggg	acccaaggct	cagtcccttt	960
ataatagctg	ttgagttcac	ggagttagta	caattatgca	tccccactga	ggctcaagag	1020
gtgaagtcac	atagccagtt	agtggtagag	atggggcttg	aacctgggtc	ttcctgactc	1080

actcagctac	ctttccatag	aaatagggac	tggaggccgg	gcgcagtggc	tcattgccttt	1140
aatccccagc	acgttggaag	gccgaagagg	gtggatcact	tgagggcagg	agttcaaaca	1200
aaaaaattaa	aaaaaaaaaa	aaaaagggcg	gcc			1233

<210> 20

<211> 1090

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (4)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (17)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (47)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (1033)

<223> n equals a,t,g, or c

<400> 20

gatntggcga	tacattnaca	cagaacagta	gacatgatac	gccaagntta	atagactact	60
atagggaaaag	ctgtacgctg	cagtaccgtc	cggaattccc	gggtcgaccc	acgcgtccgc	120
gcggtctatg	cccccagtat	cccgggtccag	ctattccgag	gacatcgtag	gctctcggag	180
aaggcgacgc	agctcctcgg	ggagcccacc	atccccgcag	agcagatggt	cctcttggga	240
tggctgttcc	cgctctcact	cccgcggccg	tgaggggccmc	aggcctcctt	ggagttagtt	300
ggacgtgggc	gctctttacc	ccttttagtcg	ctctgggtcg	cgagggcgcc	tcccaagatt	360
cgcgaactac	gccttcgcgt	cctcctggtc	gacctcgtag	agtggatatc	gctaccatcg	420
tgcactgcta	tgcagaagaa	cggcagtcag	cggaagacta	cgagaaggaa	gagagccatc	480
ggcagaggag	gctgaaggag	agagagagga	ttgggggaatt	gggagcgcc	gaagtgtggg	540
ggccgtctcc	aaagtccct	cagctagatt	ctgacgaaca	tacccagtt	gaggatgaag	600
aagaggtaac	gcatcagaaa	agcagcagtt	cagattccaa	ctcggaagaa	cataggaaam	660
agaagaccag	tcgttcaaga	aaacaagaaa	aaagaaagaa	taagtcgtct	aaaagaaagc	720
ataggaaata	ttctgatagt	gacagtaact	cagagtctga	cacaaattct	gactctgatg	780
atgataaaaa	gagagttaaa	gccaagaaga	aaaagaagaa	aaagaaacac	aaaacaaara	840
aaaagaagaa	taagaaaacc	aaaaaagaat	ccagtgactc	aagctgtaaa	gactcagaag	900
aggacttgtc	agaagctacc	tgggatggag	cagccaaatg	tggcagatac	tatggattta	960
atagggccag	aagcacctat	taatacatat	ctcttcaaga	tgaaaaacct	ttgaagtatg	1020
ggccatgctt	tgnttcccgg	tgggaaggtgc	agctatggct	gagtatgtta	aaagctggga	1080
agcgattccc						1090

<210> 21

<211> 682

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (624)

<223> n equals a,t,g, or c

<400> 21

gcaaatatta	attgccattt	actttgaaac	ctaaaatggt	caagattcca	ttttcctcca	60
ggctaataaa	taataatttg	caatatatag	attgtatttt	gtctttgaaa	cgctgtgagg	120
agatcctctt	aatgtggcat	ggtctgcttc	tatgccttgc	ttctgtgttt	cttgagctcc	180
gtggagatag	gccccctctc	ctggcttctc	tgcttgagcc	acataaaatg	ccacttcaca	240
gctcttccct	ttgaagcctg	atccagtatg	catttggagc	taattactgc	agttgacaca	300
actccatcta	aaagcgtcat	gaaagattct	gtaatcactg	ataagaaaat	gatcttgcaa	360
attattgctg	tgctctcctt	tattgcctct	ttaccttaac	agtacagttt	acaataatgt	420
aaattttttt	ctaattctttc	aactttaacc	ctagaaattg	tagatgtttt	agcagtgggt	480
atgtgatatt	ggcacaacat	aactatataa	tttgcctaat	attgtggtgc	atacctgtaa	540
tcccagctgc	tcaggagtct	gaggcatgag	aatcacatga	acccaggaga	tgagggttgc	600
gttgagctga	gagcgagtca	ctgnactcca	gccaggacga	cagagtgaaa	ccctgtctca	660
aaaaaaaaaa	aaaaaactcg	ag				682

<210> 22

<211> 770

<212> DNA

<213> Homo sapiens

<400> 22

ccaatactcc	tttactttct	ttgagttaca	atgttgcate	tattctgttc	acagccccta	60
ggctcctctt	tectgetgat	ctttctaggt	cttgactctc	tgccctcgtg	cttgaccgct	120
acccggcttc	agagtccaat	aattatattt	tcaactttgt	cctgtatatg	ctccacttct	180
tggttggaac	tctgttcagt	ttatttctctg	actttgaact	atctccacgt	agtgccacct	240
tgtttctctga	tctaaggact	gtgcaactcc	tctccagtcg	gccacatctc	tgaccaaggg	300
attcagtgaa	aactctgggt	tctcacctga	actttgatct	ttaaccccag	ctgacactag	360
taatggcctt	tgtgatcagg	ctctgaaaga	agactgactt	catgtgaacc	atgtgagcat	420
attgttcata	tatccctcaa	ggtggatcct	tcttttctaa	aaggcatcta	aaaagcaacg	480
gaagttcttt	tgaaaatcag	aggctgcctt	tttggtagca	gttctttcat	ttattctgta	540
gaggatccag	attgagctct	ttataaaata	ctctcctaca	taatgtactg	ggatagtcct	600
aacaatagta	aaccttgatc	cacagatcac	atgtgccatt	ggataaaaaat	aaataatgca	660
ragraactca	ataagaccag	cctgttttaa	ggagacatca	taaaaacctc	cattaaaaaa	720
aaaaaaaaaa	aactcgaggg	ggggcccgtg	cccaatcgcc	tgtgatgatc		770

<210> 23

<211> 565

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (10)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (21)

<223> n equals a,t,g, or c

<220>

<221> SITE
<222> (538)
<223> n equals a,t,g, or c

<400> 23

tccccccggn	ctggcaggaa	nttcgscacg	agcagaaagc	aaatcagtgg	ttttctggag	60
tttgacatgg	gggtactaac	aagggaactt	tttgggggtg	tgggaatgct	gtatatattg	120
attgtgggga	tgggtacatg	gttggatgca	tttgtcaaaa	cacacttaat	ggtaaatgcaa	180
aatgaatata	ttttatttta	tgtaaattat	acctcaaagt	tgaatttttt	taaaaagttt	240
cttttaaaaa	gtaaagacat	ttgtgggtgct	tcttgtaaat	tttactgctg	attatgacct	300
tgtttcttta	gatttgactt	tccaagtttg	aaaagaccag	tttaaaaatg	acttttgctg	360
ggcatgggtg	catgtgcacg	tagtcttacc	tactcaggag	gctgacgcag	gagggtccct	420
tgagcccagg	agttggaggc	tacagtgagc	tatgattatg	ccactgcact	ccagcttggg	480
tgaactgtg	agaccctgtc	tctaaaaaca	aaaaaaaaaa	aaaaaaactc	gaggggggnc	540
cggtaaccaa	ttcgcccaaa	atgga				565

<210> 24

<211> 1356

<212> DNA

<213> Homo sapiens

<400> 24

ggcattaag	tcttagtctc	attaattgtt	ttcaagcttt	tgcctacat	tttagactaa	60
ccctgcttat	tctgtgaat	caagtagtga	tctcctgcag	cttgggaagaa	aaaataaggg	120
atgggtaatg	taaaaatctg	gatcaatata	ctgggtcttg	gcaattatcc	tgcaaatctt	180
gccaggtaat	aaaagtgagt	aggggtgcca	taatccggaa	gtttctttgt	ttaggaaaaat	240
aaaaacaagg	aacttcatag	accccccaa	aggggaattc	tatatcttga	caagtaaaat	300
tttagatgga	aattatctac	aacaccacac	ttacggaaat	tgctatcctc	actctattat	360
ttgcaaaaagg	gttatacaca	gtagcacctt	ctaactgaag	tattaaacag	agtttccatt	420
gctgtcgtat	tttgcttaat	tattatcctt	atagcaggga	taatagttac	taacaaaaag	480
gaagcatgaa	agttttacta	tactgagcc	tggtaggact	ttttattggg	tttagtgatg	540
cagtttttaa	tgaaacatgc	cgcttttgga	ttaataacct	tagtaaagga	aatttacaga	600
tacttaaaaa	tcaaatccaa	attattgata	ggctcaggaa	aatgccagct	tcagcctgag	660
tggctacaaa	cctctttaat	aaattccagt	cttcttatgg	attggttaac	gccttattaa	720
gccctctctt	gcttatatgt	cttgatttga	tatttggact	ctgtatatac	aatactataa	780
ctcgaattgt	ttcttctcgc	ctagaagcaa	tcaaactcca	aatgggtgctg	taaactgaac	840
cacacatgga	caggccatc	ttccgaggac	ccttagattg	atcccagggg	agccctagct	900
gctattcccc	attcacgccc	ttttcagcag	gaagtagcca	gaaggagtcg	ccgccccaaa	960
tcccctaaca	gcagttagtg	tggcatctcc	acaggaagta	atggtgtagg	agttactaag	1020
aaattatttt	aggcagatag	agaggaaaag	gggtccttgg	gaagttttca	ttttttaaag	1080
catctctgga	aaagtttctt	gtaaagcccc	ggctcttaga	gccaggctgg	caacctttga	1140
tatgcaaatg	taagccatta	gaaaccaggt	ccaccagggc	cagggtgtgt	gctcacgcct	1200
gtaatcccaa	cactttggga	agcctaggca	ggtggatcac	ctgaggtcag	gagttcgaga	1260
ccagcctggc	caacatggtg	aaaccccgcc	tctaataaaa	acacaaaaaa	ctaaaaaaa	1320
aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaa			1356

<210> 25

<211> 617

<212> DNA

<213> Homo sapiens

<400> 25

ggcacagcac	agcctgagat	cttggggatc	cctcagccta	acaccacag	acgtcagctg	60
gtggattccc	gctgcatcaa	ggcctaccca	ctgtctccat	gctgggctct	ccctgccttc	120
tgtggctcct	ggcgtgacc	ttcttggttc	ccagagctca	gcccttgccc	cctcaagact	180

ttgaagaaga	ggaggcagat	gagactgaga	cggcgtggcc	gcctttgccg	gctgtccccc	240
gcgactacga	ccactgccga	cacctgcagg	tgccctgcaa	ggagctacag	agggtcgggc	300
cggcggcctg	cctgtgcca	ggaytctcca	gccccgcca	gccgcccac	ccgcccgcga	360
tgggagaagt	gcgcattgcg	gccgaagagg	gccgcgcagt	ggtccactgg	tgtgccccct	420
tctccccgt	cctccactac	tggctgctgc	tttgggacgg	cagcgagstg	cgcagaaggg	480
gccccgcgct	gaacgctacg	gtccgcagag	ccgaactgaa	ggggctgaag	ccagggggca	540
tttatgtcgt	ttgcgtagt	gccgctaacg	aggccggggc	aagccgcgtg	ccccaggctg	600
gaggagagg	cctcgag					617

<210> 26
 <211> 648
 <212> DNA
 <213> Homo sapiens

<400> 26	
ggcacgaggt	gttttaaacc
gatgactatt	catgctctgc
ttccaccta	catttctgtt
tttttcttta	aaagataggt
ataaccattt	cggtagagaa
taaaaatata	attttttaaa
gaaaatggca	aatacttttc
cctttagtga	aaggtacac
gccaccttct	taaagctcag
tggggcatta	aaatcccatg
tgaaatgtcc	atcttagttt
	tgtaaaaaaa
	aaaaaaaaaa
	aaaaaaaaaa
	60
	120
	180
	240
	300
	360
	420
	480
	540
	600
	648

<210> 27
 <211> 1388
 <212> DNA
 <213> Homo sapiens

<400> 27	
ggcacgaggt	aagttgcaag
actaatgaag	accgcatcta
catcatactt	agattttatat
tgagatacca	tttggcctca
gactagtagg	aaaagaaggc
tacataaaat	gagtaataac
tgcttaatat	taacatttta
gactcatgag	gacattgggtg
gaggtgaact	tctttacaca
actgtggcta	caagaatctc
cagacactga	actaatactt
attttccaaa	gaaacagagg
aggcattcta	gctctcacat
cttctcattg	tcacacttgg
ggtgttacca	atttactaga
attctgtggg	ggtaaaacca
aatctgttat	tttgaattgt
acaaaccatc	gagcctcatc
aggatatagtg	ccgggctggg
gggtggatca	cctgaggtca
ctttactaaa	aatacaaaat
caggaggctg	aggcaggaga
	atcgcttgaa
	cccaggaggc
	ggaggttggtg
	gtgagccgag
	60
	120
	180
	240
	300
	360
	420
	480
	540
	600
	660
	720
	780
	840
	900
	960
	1020
	1080
	1140
	1200
	1260
	1320

atcacgtcac tgcactccag cctgggcaac aagagtgaac ttccatctca aaaaaaaaaa 1380
 aaaaaaaaaa 1388

<210> 28
 <211> 616
 <212> DNA
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (17)
 <223> n equals a,t,g, or c

<220>
 <221> SITE
 <222> (580)
 <223> n equals a,t,g, or c

<400> 28
 cmgctrctra gcaactnagt gggatcccc gggctgcagg aattcggcac gaggagaacg 60
 gctgcacgtg ggagatgctc cgtggatgtt tgtagaacgc tggcttccgt gtttcctcgt 120
 tgtggctgtg gtgggtgtggg tctttgcctg tggacccgtg gaagacaaag aagacagttt 180
 tggatgggtca agctattttc ttgcttcagg gctccctccc ctgctttttg aagcctcaca 240
 aaccaggact gtgagggcag gaaggcttgg ggtctttgtg tgctgagcct cattaggggt 300
 ttaagaacct ccttcctttc atctctagct tacgagaggg atgattcatt atcttcctc 360
 ctgaggtgca agtagaagca gacagtctct gcctccctgc ttgcctttcc tccctcccat 420
 tcaactgttga ttattgccct caagaataac aggttgccca gctactcgag argcttaagt 480
 gggaggattg cttgacccca ggagttcgag gctgcagtga gctatgatcg cttcactgag 540
 ctatagcctg gcagacacag agagacccta tctcaagcan acagacaaac aaaaaaaaaa 600
 aaaaaaaaaa ctcgag 616

<210> 29
 <211> 828
 <212> DNA
 <213> Homo sapiens

<400> 29
 acgagaacac catgctagt agttcattcc taacagagga gaacttgcat cttgactaag 60
 cattagtgat ctcaaatcct ctgcttatga tttttaaact tctgatcttc agaataatttt 120
 tccatgagct agctctggct ttgtgcatct caaaccttgt ttctctccca tggctgtcat 180
 acttctggtg ccctgagatg cagaatttat ttctacttga tacacacatt tgggtattga 240
 tgtagggtta gtacagcagg taggttgaga atttctggag cctccctccc tccctttgtt 300
 ctgacctttc cttagtcata tcatcctaga aagatcttcc ctggcttccg ttaaaacatg 360
 gcctctcatt tcattctctc cctgacaacc ctgatgtagt ttccatttca ggactcatca 420
 ccccaacatc ctttctctgt tacagcccat ctccctgctc agaacacaag ctctgagagg 480
 tggaaaggcct ctattgtggg ttttggcgaa tccccaatct ctagatgggtg tctggcatgt 540
 gatagagatt caacaaacac ttcaacaaat aatgaataaa gttaaatttt tcagagtgc 600
 atcatgcctc tcccttctct tgccagggcg gaggtctgtc ctgggtttgag cggcttctgc 660
 agctccagct ccttgtaact agtctggaga atgatggagc tcagtccatt ttaatcccat 720
 gaacattaaa tgcgtggatg tgtggatgct gggatggatg gatgacgctc ctagcacggc 780
 agcttgcagg ggattggcga tttccagtaa ggtgtgctaa gactcgag 828

<210> 30
 <211> 581

<212> DNA

<213> Homo sapiens

<400> 30

ggcacgagat	tatgggggca	gtttcccca	tgctgttctt	gtgatagtag	ggaattctca	60
tgagagctga	tggttttaag	tgtggcactt	cttcacgctc	tctctcacct	gatgccatgt	120
aagacgtgcc	ttgcttccac	ttcaccttct	gccatgattg	taagtttcct	gaggcctccc	180
cagccagcca	tgtggaactg	tgagtcaatt	aaaccttttt	tgtttataca	ttaccagtc	240
tcaggtagta	tctttatagc	agtgtgagaa	tggactaatg	cagtctcttt	gttagattgc	300
ctatgtttct	tgactgtcta	ataagatatg	actttcagtg	atacaaatga	tcagagaacc	360
ttgcaaaagc	gatgtggtgg	agagagttca	gggagaacca	gagtggaaaca	aaccggagga	420
atgaatatgg	tagagtgcaa	gtgaatttgc	tcctgatttc	cagaagcagt	tgtagcaaac	480
agagtttacc	cttgtatagc	attaaaatag	tctaggacaa	accagaaagt	ccataatgtc	540
ctccagcatt	aagagagcac	ctcgtgccga	attcggcacg	a		581

<210> 31

<211> 789

<212> DNA

<213> Homo sapiens

<400> 31

ggcacgagcc	tccttccaga	agttccaggg	ttttctttga	tggttgccat	tctgcttaga	60
gaacttccat	tagcctttct	tttgggtggg	tcttctggtg	acaaattctg	tttcaacttc	120
tctgagaatg	ttttgctttc	cttttcatcc	ctgaaggaca	tttttgctgg	atataagaat	180
tctgggttaa	tggttctttt	cattgtttta	aaaatatatt	gtactttcag	ctgggctcca	240
tggtttctga	tgagaaattc	gctgtcattt	gacttgttta	tgacagatag	ttaaggcagt	300
ttttgttttag	tggcttttga	aatgttttgt	cttttgtttt	ttggagtttg	attattgtat	360
gtcttagttt	ggatttcttt	gggttcaccc	tgtttagggg	ttgcttacct	aagatctgta	420
gatttatgtc	tcttgccaaa	tttgggaact	tttaagccat	cattcgtaga	gtacagtttc	480
aacccacact	tctttctcct	gtcccttcgt	gagttcagtg	actggagttg	ttatagtcct	540
ataggtcccc	aagactgggt	tttttttttt	tttttttttt	ttttgctggg	acatttcttt	600
ctctcttccc	ttcccatccc	gtcccatcca	gtcccgctcc	gtcccatccc	ttcctgtccc	660
ttcccatccc	ttcccgctcc	ttcccatctt	cggagtctct	ccctgtttcc	caggctggag	720
tgcaatggac	gttctcggct	cactgcaacc	gccgcctccc	tagttcgaat	gattcttctg	780
tctcagcct						789

<210> 32

<211> 884

<212> DNA

<213> Homo sapiens

<400> 32

ggcacgagca	gaatcagggc	actgagctct	actgtaagtg	tgtatttaat	ctccattttt	60
aatttgaccc	atagaaaatt	gtcttttatag	accaaaaatg	gtcaaagatg	agcaatttca	120
tatctattac	atgcttagtg	ttcactatct	tggggcatct	tgtttctctc	caggttgctc	180
attcatctgt	ttttgagttt	aaaaccttgt	atgtgcttaa	gaccaataga	tattctcagt	240
cacttttttag	acatttttgt	cacctcagtt	ttatcagaac	tagaaaaatc	ttccttaaaa	300
acaactgaaa	ccttttcttt	tgcagtgttt	cttttgctaa	tgatgaacta	atattaacaa	360
cttaccttct	aataactttg	tctttgtaac	tcaggtttaa	aagcattact	ccaccaatct	420
cttttattct	ccattaaaag	atattatttc	ctttataatt	catttttatt	ccttatttgt	480
ttagtgcctt	aagtctgttt	ctgtacagct	tggagactgt	aacttggaat	aatgtagata	540
tattcattat	atactatctc	acttaaatgt	gaattctgaa	gaagtttttc	tgagaaaata	600
aattcgtctt	gtttgagtg	tccttctctg	tctgctgtat	tctaccaggg	ggacagagga	660
aagaagggca	gagaaattct	ttcctggatt	tcatgaaatc	actgaggcca	gagcaacaga	720
tcattttaact	tttctttatc	ccgtgtatgt	attactgaaa	aaaatgggtac	ataatagaat	780

ttgagtttat	ttttatctga	cttcaggtac	atagggacac	ttcctgttta	aaagaactct	840
gcattgaatt	gtgcctttca	taaagcaaaa	aaaaaaaaaa	aaaa		884

<210> 33
 <211> 866
 <212> DNA
 <213> Homo sapiens

<400> 33						
ccacgcgtcc	gctccaaaca	aacaaaaaat	gaactttatt	tagatatatt	ttacatatga	60
tgaagtattt	ttttgatgta	gtagtttttc	tcaccttctt	tttagtcttc	tctttatcca	120
tttttctttc	tgatgaagaa	ttccctgtga	gtaggaccca	gaacataggg	ctttgtcatt	180
tcaacccttc	gttctctgaa	taggctgttt	attggcaaca	ttactggaa	acattttatg	240
tacagcattg	gagtcctact	ctgtcgcttc	agctcactgc	aacctccgcc	tcctgggttc	300
aagtgatgtg	cactgtatga	actgtgagag	caagcatatc	attataacat	tggacaatga	360
gccaaagacag	ttctgatgga	cttttgaaga	gggatttttc	aaaagcattt	aactcatcat	420
attaataaaa	taaatcctat	gatttatggg	aaattctgtr	ggatcaactt	tggaaactgt	480
ttactataaa	ggtagcatgc	gtaggcatga	atcttgataa	gacaagattc	tgatccgggg	540
ttctgagtgg	gtccttatat	tctgcagagc	tgaaccaggt	ggaataggag	gagagtttgg	600
gtaacagtca	aacacaacat	ccaaaattat	gttgaatgca	gtggtgagag	ctattccctt	660
taaaactctc	tcttggttct	tctgactgtg	tcaagaatac	tgtatttggt	tgggtactgg	720
ctgggttttt	tttttttttt	tttgaaatgc	actccagcct	gggcgacaag	agtgaaactc	780
tgtctgaaaa	gaaagaaaga	aagaaaaaga	aagaaaggaa	agaagggaag	aagaaaaaaa	840
agaaaagaaa	gaaaaaaaaa	aaaaaa				866

<210> 34
 <211> 1694
 <212> DNA
 <213> Homo sapiens

<400> 34						
ccacgcgtcc	gggataaact	atztatcata	aattttaaag	ttgttattgt	gcctttattg	60
cacatttttt	ctcatgcctt	ttattataaa	atataacttg	tttcaccttg	tttttggtca	120
aatcccagtt	actgtccatg	taaacatatg	gcaacataag	aacgtcactt	tttttatcct	180
gcactgtggc	atacctgctc	ttacaagaga	ttctgttgca	cttacaatatt	caaagtatgg	240
cacagttata	gagactttat	tattcttaat	tctttattta	gacttaaaca	ttatttggtg	300
ctaattaaaa	cagttcatat	gatgaacttt	ataataaaa	atataatttt	caggccaggc	360
atagtggctc	atgtctgtaa	tcccagcact	ccaggagtct	gaggcaggcg	gaccacttga	420
gcccagagtg	ttgagaccag	cctggacaac	atagggaggt	cctgtctcca	caaaaaattt	480
ttaaaaaatt	agctaggcat	tgtggtgcat	tcttgtggtc	ccagctactc	cggaggctga	540
ggtaggagga	ccacttgagc	cccagaggtc	aaggcgagc	taagttgtga	gagtgagact	600
ctgtctcaat	gataataatc	ataatcataa	taagatatg	tacccagag	tcaagaatgc	660
aataagctgt	gagagtga	ccctgtctca	atgatcttaa	taatagtcac	aataaagata	720
tatatatttg	atcagtctgt	gggtaattga	atgatgattg	ctggagttaa	actgagcaga	780
atgggtacagt	gctgtactgt	acagtacata	agaatcagga	gaggcagctg	ttttaagccc	840
cagtcttatt	atgcccattt	aatcaaggca	tgtgtgctcc	tgcatgaat	tattctgatg	900
ttgagtaaaa	ctgtcatctt	caattcatat	cactttcaaa	tgttacatta	taagagaaat	960
gtaacctggg	aagtactaag	ttgtagattg	ttacagagtc	agaaaagttg	gatgtaattg	1020
ctgatattga	ttatggtaag	tagtctttcc	ttgctaaagg	aatgccaat	tttaaacaaa	1080
ttataaagca	atcattttta	cacaaagatg	aaaatacaca	tgggtcctca	tatgctaaaa	1140
tgttatatgt	ataggtttta	atatctaatt	tggaaaagct	agacataata	atagggtttc	1200
atccttgag	ctaccaatat	caaactatac	agttcttggt	agtacctcct	ctagggtggt	1260
tggagagggg	gcagagaaaa	gagctcaaat	tcaagttggt	tttatatctg	acataatttt	1320
ttaaaattat	aagtcaaatg	tttgttaagg	gaggggattt	aagttattaa	ttaaaataag	1380
attttttagat	cttcagattt	tacttatcta	tctaccttc	cattgttatt	gagagtgtgc	1440

tgtattttct	ttttttttta	ctccctggcc	tgattttaaag	atactctgga	acattccaga	1500
ggcactgaat	ttatcattct	aaaagggcca	tgtagtactc	ccaggaggca	tagagctgag	1560
ccactcaatg	tgtcttgagg	ttccaaaatt	ctgattctat	gcttttatga	acatgtaatt	1620
tagcagatgt	tacctaatag	caaaagaaaa	tgctatcttt	accagaatgc	acttaggaaa	1680
aaaaaaaaaa	aaaa					1694

<210> 35

<211> 1215

<212> DNA

<213> Homo sapiens

<400> 35

ggcagcagct	tttcagtact	ctcacttatt	catgacacag	gaatgaccct	ttactcaaaa	60
ctcttggtgt	tggtcaaagg	tgagcttctt	ttcccttag	tcttagccta	tggtgtgctg	120
ttgtatatgt	ttaccaagtt	caactaccta	attttgaaag	tctttccaaa	taagatacaa	180
attaaaaggg	gaagcattgc	cagtaacagg	tccctagaga	gcagtgccag	cctgcctgca	240
agaaaagagg	agaaacttct	taaaaagttt	taagcctggg	caacataagg	agattgtttc	300
tatgaaaaat	aaaaattagc	caggatgggtg	gtgtacacct	gtagtcccag	ctactcggga	360
agatgagatt	ggaggatcac	ttgggccttg	gaggttgagg	ctacagtga	ctgtgattgt	420
gccactgtac	tccagtctgg	gcgacagtga	aatcttgtct	aaacaaaaac	aatttaactg	480
ggaagcacag	tggtccttga	ggacatttaa	tatcaggaca	aagagcctat	gaatataatca	540
ctgatgtata	taaaccctaa	ggcgtaata	aaagctaact	gtttagtgtt	atccatttaa	600
gggaacagga	ggaattgcat	aacttttaga	ttagtcatag	tggtgcctaa	gggatatgct	660
gtggtatatt	tgtatagcca	gggcacttag	ccttccaacc	aatttatata	ccatgttctt	720
caactgtggg	tgagatttag	cctcaagatt	tgatttacta	tatgtaagta	cattacttga	780
tttctataaa	gaatcttttag	tggaagatgt	tattctgaat	tatttatcaa	tatgattaat	840
ccagttagaa	attattaatg	atcttccttt	atactataca	taggataact	tttaacttgt	900
cgctacagtt	gttgctctga	ggatcttaat	tttgttactt	tctaggctac	atgaagctat	960
ctttttaaaa	agtgtactct	tcattttcac	tggtattgtg	tacttagcat	aaaaaactca	1020
aatctaggcc	aggtgcagtg	gctcatgcct	gtaatcccag	cactttggga	ggctgaagca	1080
tgcggtacac	ttgagcccag	gagttcaaga	ccagcctggg	caacatggct	aatgaacact	1140
aagcatagtt	ttgatgtctt	tcataattaa	gactttctca	aattctttaa	aaaaaaaaaa	1200
aaaaaaaaaa	aaaaa					1215

<210> 36

<211> 1794

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (1675)

<223> n equals a,t,g, or c

<400> 36

gctcctctag	tgactgctgg	ggtgctagtt	ttcacggtta	ccacagaact	tgagagagga	60
aaatgagaac	ggcaaattaa	aatgccataa	aattccattc	ttaacaagcg	ttagtttttt	120
tttccctatat	atccactccc	tcacctatct	tatgacatat	gagaaatcct	aacatgggac	180
ttggtgtata	attagcattg	aatgaattta	agttttcttt	ctttctttct	tttcttttat	240
cttctgtggt	ctcctgcaga	atcagtctat	aaaaagggca	tggtaaaaaa	aaatctatct	300
catagcattg	ttgaaaagat	taaatgacat	aataagatga	tgcatatata	gtagctagca	360
ctgtacctga	tgcatattag	gagcttgata	atattactaa	cattatcatc	atcaatgcta	420
tctgagcaaa	gaggctgttc	cttctttcca	gccagagttc	ttctgttgga	taatatattcc	480
agctgtgaac	cccaaccaga	gaccttgaa	cctttatttc	ctttctctct	attggccttt	540
ggctgaagtc	tcctttctgc	agagaacaaa	gtgaccagct	tttgatgaac	tattttcctc	600

ttttatccat	ttccaagtgt	tagccatagt	gaagaagtgt	agctgggtgc	tagcccaact	660
caagaagtga	taatgtwata	tccaacccaa	gataaaactca	aggataaactt	tcaacacggc	720
tactcaagca	gttcaggggg	gaggcatctt	gtagagagga	gaccaggaag	tcacttggca	780
gctgggcccag	cacacagagg	gccattgctc	cagtaaagca	gctaacctcc	atctcttcat	840
caaaccatcc	tagactcagc	actctgcaga	gaaggagcag	atggagggaa	tgtgtggaga	900
gattagataa	gaaggatttc	taatctagt	ggggagcgag	taaaatgtac	agaagtttga	960
gaagccaagc	tcttatgtam	maatccrccc	ccatcactca	acaatcccca	ctgtatgaat	1020
aaagcctgga	aagttttccaa	ttaaaacagt	gcttgtgaat	attggcaagg	ggtatctttg	1080
tgtgcagggt	acattttaaag	ggagaagggt	gaagatacac	ccttgccctty	taggagtaca	1140
ctttctgaga	gcttgttcac	caggctgtga	gtttctcagt	ctatctgttt	ctcagctctgt	1200
tagtaatgaa	tgtatctccc	acttagcaca	gcagctagca	catagtagat	gcctaacaaa	1260
tgtgtgttaa	attgaatgtt	ggaagtcctgt	gtcctgaaaag	cttttcttca	catattacaa	1320
ggctttttat	ttgttcaaca	cattttttacc	aagttttttt	ctgtgttcca	ggtcttttga	1380
ttagctgttt	ttaagtcaca	gaaatgggtg	tcgggaacaa	aaccagtcaa	aagtcctcat	1440
tctacatcat	ttacactttc	ccttcctata	tttataagtt	ttaatatcag	cttctacaat	1500
aggttccaga	acaagtgggtg	ctcaaggaat	ggagaaaatg	actattccaa	ccctagctgt	1560
aggtgaacca	aaaaccccag	agaaatcaaaa	gtgtagttta	aagcagtgtc	tctcaagttg	1620
taatgtgcat	atagatcacc	tgggtgtgtt	attaaaaatgc	aaattctaaa	agagncagga	1680
gaatatcttg	agcctgggag	ccagagggtg	cagtgaagccg	agatcattcc	actgcattcc	1740
agcctgggtg	acacagcgag	actccatctt	aaaaaaaaaa	aaaaaaaaact	cgta	1794

<210> 37

<211> 1174

<212> DNA

<213> Homo sapiens

<400> 37

ggcacgagca	aaggcaggaa	cttttgtcat	ttttgttctt	caagatgtcc	ccaatgccta	60
cagcagtaag	tgcactgctg	ggcacacagt	aggtattcaa	taaacattca	atgaatgaat	120
gaatgaatga	atgaatgaat	gcattgccaa	accttgcatg	gctgcctttt	gttcctgccc	180
tagctgctgc	ctccccagcc	ggcctggctg	ctccccagag	cagagatgtg	ccttttctctg	240
ttagccctgc	cacacagttg	aacattgggt	agagcccatg	ggccaggggc	agaggcagga	300
acacactcag	ggcagcgtgc	tgccttcctc	ccatccttct	agaggcaaaag	ccaccacagt	360
ccattcctgt	tgccaaagagc	cttgggggta	ggggaaagtag	acaggaattc	accactcctt	420
aaaaccagag	aagtccagag	tcctctgagg	gacagggtc	ttgatgttca	gagggaaatca	480
gcctcggcct	ggggaggctg	ggatttgggt	ggtttattct	tcagcatcca	cttcttctcc	540
agggagtctg	tcagtgcag	gaaaagtgc	tcagaacccc	agagccaggc	ctagtggaca	600
ctaggttctg	cctgtcccat	ggagggtgct	gtggctttga	cggattagaa	gtcaggactt	660
ggccatcatg	actacatgct	ttaaaaaata	ttagttttct	ttttttgttt	tgggtttttt	720
aaggcggtgg	ggggctactt	tgtgtttagg	ttttacatcc	tttgcaataa	agttccaccc	780
atccagcttg	tgcagtgaag	aagagggaag	ggacttcagt	ggctttgcct	tgtctataca	840
tgggcccagag	agaaaaaagg	aagagggtg	ggcgcggtg	ctcacgcctg	tgggtcccagt	900
actttgggag	gccgaggtg	gaggatcacc	tgaggtcagt	agttgagacc	agcctggcca	960
acatggcaaa	atcccgctctc	tactaaaaat	acaaaaatta	cctgggcgtg	gtggtgggca	1020
tctgtaatcc	cagctactcc	agaggctgag	gcgggagaat	ggcttgaacc	caggagcat	1080
aggctgccag	tgagccgaga	tcatcccatg	gcactctagc	ctaagggata	gagtgagact	1140
ctgtctcaaa	aaaaaaaaaa	aaaaaaaaaa	aaaa			1174

<210> 38

<211> 1087

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (408)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (1005)

<223> n equals a,t,g, or c

<400> 38

gtccattctt accaccactc tcccaggtaa atgctccatc ttctctgtct tgggttgcat	60
ttggtccac ctggctctct ttcagttaac tcccttcagt ccaccaatg cagtcttttc	120
tctgcagcca aatttttttc tatagttcag gtctgatgat accactttcc attttcaaac	180
tccttaaatt ggcctcggag acctgtgatg gccaatcttc cccaccctcg tgggtcagggt	240
tcctctggtg tgggcctatc tcttagtgca aggcctccctg tagactgtcc tggggctgcc	300
tgtctaattg tcttctcatg aaccaggtag gcargcagga accactcatc tagtctgggg	360
gtgactgggg ggtgcgaaag ttttgttaca ggatgtattt aacacttnca ttgtctgggt	420
caaaattctc tttgaagaaa ctgcctttgt ttaattctgt ttawttattt agaacttacc	480
tagaattgat ttgcacttgc aaaaatgctg ctgttttaga gtttgcttcc taggctgggt	540
gcttgctctg tcaccttagc taaaaactag gtactggagt gccttttcta ttcycttctc	600
ttctttcttt cttcttcttt ttttttktt ttttagtggt agtctgggtt tgcgccccg	660
gcttaagtgc agtggcacga tcccggctca ctgcmctcc cgcctcccag gakaagccct	720
tgaacctggg agggagttgc ggtgagcgga gatcgggcca ttgcactcca gcccgggcaa	780
cagatcctga ctcttatcaa aaaaaaaaaa aaaaagaaat taaccttagt ttaactttct	840
tgctttacca agtttttaat ttttaaaatt ttttactctt ttgtaatagc acttagttta	900
aaacacaaca cattgtacag ctgtactaaa tatactctca tttatatcat tgttccatag	960
ctttttcttt ttgtgtgtgt gtgtgtgtgt gtatgtgtgg gtgancagtg agatccgtct	1020
caaaaaaaaa ccaaccaaac aaaaaaaaaa aaaaaaactc gtaggggggg atccggtacc	1080
caattcg	1087

<210> 39

<211> 438

<212> DNA

<213> Homo sapiens

<400> 39

ggcacgagggc aaatgccata tatgttttaga ccagcctttc tcaattgtgg cactttttgcc	60
atttttggcc agttaaatc tggtgtgggg gctgtcctgt gcattgcagg atgttttagca	120
gcattctctg cctctacctc ctaaatgccg gtgcacctc cctgcagtta aatgacccca	180
aatgtctcca gacataggca aatgttcccc ggagaacaaa gtcattcttg gttgaggaac	240
aaaatgaacc tatatttaag atggctctaaa tgtttctgag atggctctaaa catttataag	300
cttaacaatt aacttgaaac tcttcaggac tagagaagct atttgtaaat ttaaacatct	360
gacgccttga aacaccttta tcataaagat aactaaacca gtgttctttg tgaatggcca	420
aaaaaaaaa aaaaaaaa	438

<210> 40

<211> 734

<212> DNA

<213> Homo sapiens

<400> 40

gctcgtgccg ctgctgggca ctgggagcag ggggcggcca aaggcagtggt gtgggcaggt	60
ccatgcctcc cctggccccc cagctctgca gggcagtggt cctgggttcc atcttctgctc	120
tgctgcaggt gaagcctctg aacgggagcc caggcccccag agatgggagc cagacagaga	180
aaacgcccctc tgcagaccag aatcaagaac agttcgaaga gcactttgtg gcctcctcag	240
tgggtgagat gtggcaggtg gtggacatgg ccagcagga agaagaccag tcgtccaaga	300

cggcagctgt	tcacaagcac	tctttccacc	tcagcttctg	ctttagtctg	gccagtgtca	360
tggttttctc	aggaggggcca	ttgaggcgga	cattcccaaa	tatccaactc	tgcttcatgc	420
tcactcactg	accctccctc	cctcctgggc	tccagggtcac	aactcccaaa	ggagatgcag	480
gcatggctct	ctgcctctga	tcaccatcac	tgtatctcaa	ggttcagcag	cagagatacc	540
agttgccatc	agtgcctaact	gactgcctct	ccagggttcgg	agtttcatct	cccaggggcca	600
gagacagcag	accacatcc	ttctctccca	cacctctcct	ggttttgttc	aggacagcag	660
attagaggca	ggaggcaatg	acaataaaat	aacgataaaa	tcctgagaac	aaaaaaaaaa	720
aaaaaaaaact	cgag					734

<210> 41

<211> 1346

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (707)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (998)

<223> n equals a,t,g, or c

<400> 41

gggcagmgya	aacacctctt	cttataaaca	tggccctgct	ttagcaacag	gtctccatta	60
tgaagcaatt	tggatttgga	cactctatca	agttacttaa	aactaaactc	tgccgtatag	120
tgttttactt	ggtatttttt	gtgtggccac	agtctagtgt	gatcagagaa	gccacacaga	180
cataaattcc	cagtcctgat	tccattgatt	attagatatg	tacttaccct	gaacatctgt	240
aagatgggga	gtaatgatgt	ttaccacaaa	gagttagtgt	aaggattcga	tgagagaact	300
tgcatttgag	ccatctggag	agaatgttag	ttttagcact	aacgatcgtt	tataatactg	360
ctgaggagga	aggactgaca	tcaccctgct	tgagaatcta	gtatatgttg	ggcattccyt	420
atccattcat	gagcctgaga	gtacttgaga	wgttttgatc	ttggacactg	ytgtgcagct	480
ttaaggatga	tgagggggaa	atggaaagag	tcytgaaggg	accagtgtca	gagacagata	540
gaaaagggct	gcctgagccg	ggcgagtgga	ctcacgcctg	taatcccaac	actttgggag	600
gccggggcgg	gtggatcacc	tgagggtcgg	agttcgagac	ccgtctaacc	aacatggaga	660
aacctgtctt	ctactawaaa	tacaaaatta	gccagggtatg	gtggcanatg	cctgttaccc	720
cagctgcttg	ggaggctgag	gcaggagaat	tgcttgagag	gcagagggtg	ttgtaagccg	780
agwtcgcgcc	attgcacgcc	agcctgggca	acaagatkga	aactcccatc	tcaaaaaaag	840
aaaagaaaag	ggctgcctgg	gtccctggga	agaccagct	cccctctgta	tcctagccaa	900
ctccttaagt	gcacacaacc	tcagctgacc	acataggtgg	cagtgatagc	aaatggtgag	960
gtgggtgggg	gtgggggtcc	tgcagctcac	ttctgtntct	gcgtaataag	ggagagatca	1020
gagatgcata	gactttttaa	ctggtacggt	tcttagagat	ggtccttggc	cttctgttgt	1080
tgttgttgtt	tttttctttt	tcttctcttc	cttctccttc	ttcttctctt	ctccttcttt	1140
cttctttttt	ttttcagagt	cttgctctgt	caccaagact	ggagtgaagt	gatgtgatct	1200
cggcttactg	caacctccac	ctcctgaggt	caggagaatt	gcttgagagg	cagaggttgt	1260
tgtaagccga	gatcgcgcca	ttgcacgcca	gcctgggcaa	caagattgaa	actcccatct	1320
caaaaaaaaa	aaaaaaaaac	tcgtag				1346

<210> 42

<211> 998

<212> DNA

<213> Homo sapiens

<400> 42

gggtgctctgc	ccgtttttgt	ccttaatggc	cggtgtgatt	cgataggaat	cgacagatgg	60
gagccaagag	gggacgtaac	cacttgccag	tgtggcccag	tatttatatg	gtagaagctc	120
gggctggcat	tgagatggag	gtgggaaaga	ggaggggtgc	tcgtgtttaa	agattttctc	180
cttgtgtggt	ccttggttat	ggctcccat	gggtgcttgc	gctcccttgg	tcgcctttcc	240
ttgcattttg	ctttttgcct	ttagtccttc	tgcagtcagg	gaccatgttg	gtgactctcg	300
gtcagatggt	cccatatttg	catgtcttgc	tttggcctct	ctggctctag	gttctgtatt	360
gctagttgct	ttctgacatt	ttttctcatt	gtttgtgcta	ccttggccaa	gaacctcttt	420
gccttctccc	tcgtttgtaa	cattgagatc	ctagtagcag	ttacttctta	gggttgtag	480
gatgacatga	gatgggag	ttggccatct	gccacacaag	agttcccttc	tcactgccat	540
cctctgccca	gggtgttccc	cagaatctgc	agggcccat	tggatcatct	gctgtctgca	600
caccttcttc	tctcacctct	tggcacttcc	ctcaaaaaag	agagaagtgg	agcacagtaa	660
ataaacgcca	gcgttttctc	cagttcccag	cacctctccg	gaactggatc	ccccaaactc	720
ctctgtcact	ttctgtgtcc	agcgggcccc	tgggtctctt	cactgtcttc	actctgtctc	780
gcctgtgctc	ttggccttgg	tgctgtaggg	actgttataa	gagctgctgt	ccgatcccca	840
tattcaatct	cacagcccca	ctttgtgtac	acacaccaga	gccatcttct	taaaatcaga	900
tcctgtcact	cctctcccta	aaatctacca	gtggtcta	aaaatcttga	atttttagca	960
taaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaactcga			998

<210> 43

<211> 658

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (6)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (15)

<223> n equals a,t,g, or c

<400> 43

acgacnagac	aggtnaccgg	taccggaatt	acccgggtac	gaaccacgc	gtaccgatat	60
ttcgtcattt	aagaatatct	acagatatgc	ctttgatttt	gcaagggata	aagatcagag	120
aagccttgat	attgatactg	ctaaatctat	gttagctctt	ctgcttgga	ggacatggcc	180
actgttttca	gtattttacc	agtacctgga	gcaatcaaag	tatcgtgtta	tgaacaaaga	240
tcaatggtag	aatgtattag	aattcagcag	aacagtcctt	gctgatctta	gtaactatga	300
tgaagatggt	gcttggcctg	ttcttcttga	tgaatttgtt	gagtggcaaa	aagtccgtca	360
gacatcatag	caagaactat	gtgaagaaaa	tgcaaacctt	tcaattccca	cgtgtatata	420
agctaagtgt	atgaggggga	aaaaaatcca	acgggtgcat	tttcattcat	atgaaagact	480
tctcatagta	cttttttttc	ctttttttta	aggaggtttt	tcttgttaca	tgtgatgggc	540
attgagccac	acctcttctt	agactgaata	ttgaagtgtt	tgttttgagt	tatgtttata	600
acatttattt	cagaacaata	aagattcaga	tttgtgacaa	aaaaaaaaaa	aaaaaaaaaa	658

<210> 44

<211> 566

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (68)

<223> n equals a,t,g, or c

<400> 44

ggctatttag	gtgccctata	gggaaagctg	gtacgcctgc	aggtmccggt	ccggaattcc	60
cgggtcgncc	cacgcgtccg	gtcagagaga	aagaactgac	tgaaacgttt	gagatgaaga	120
aagttctcct	cctgatcaca	gccatcttgg	cagtggctgt	tggtttccca	gtctctcaag	180
accaggaacg	agaaaaaaga	agtatcagt	acagcgatga	attagcttca	gggttttttg	240
tgttccctta	cccatatcca	tttcgcccac	ttccaccaat	tccattttcca	agatttccat	300
ggtttagacg	taatttttct	attccaatac	ctgaatctgc	ccctacaact	ccccttcccta	360
gcgaaaagta	aacaagaagg	aaaagtcacg	ataaacctgg	tcacctgaaa	ttgaaattga	420
gccacttctt	tgaagaatca	aaattcctgt	taataaaaaga	aaaacaaatg	taattgaaat	480
agcacacagc	attctctagt	caatatcttt	agtgatcttc	tttaataaac	atgaaagcaa	540
aaaaaaaaaa	aaaaaagggg	ggccgc				566

<210> 45

<211> 1277

<212> DNA

<213> Homo sapiens

<400> 45

ggcacgagga	ataatcctag	ctctatctgc	aggggtttat	cgtgggtttcc	ctgagtttaa	60
acatgtaaca	ggggagcaga	aggagggata	aatgatttgt	ttatgctcaa	taaagatgtt	120
gctgctgttc	tgtagcctaa	cctttgctct	cataacctgc	ataaacttgc	aaagtcttta	180
tttattctcc	taccagcaaa	ttattggaat	tcatagtcat	gtctaatttt	agttagctta	240
gggatggcac	tcaacagggt	tataaaagg	ggagaacatc	tctctcattc	ctgaatttcc	300
aaggctggga	ttacggtaaa	gctgtgggag	atcacagtaa	ttgagggttc	ccaggacact	360
cccatagtgg	gttatgcttc	ccaaatctgg	ctttctaagc	ataatggaat	ctggcctaac	420
ctgtcttcca	atatttgagt	ctcttctata	gcaccctgt	cagagtgggt	tgtgttcttc	480
acttgataaa	aacttttttg	gggatgcac	attacagttt	gagccagggt	atttcatctt	540
caggcagctg	tgatagataa	taatacacaa	taaggaaaaca	acgaaacact	ttttaatgga	600
ttagtaagt	ggaatagata	gggcacctgt	catttcattg	ttgtccgcat	gacctgatga	660
ggtagttaac	tgtagtgct	acttccattt	tcaggccaaa	aactgggggc	ttggaatagt	720
ggagcaactc	cggcagtgcc	acacagctat	gaaatatcag	gtcagaacct	cagggaact	780
ggatccaaaa	acctgagagg	aagctatata	tcctttaact	tcttcaccca	aagctgaaat	840
ctgcttcccta	gttctatcct	ccaatgacct	aaggaaaaaga	atctcttccg	ctactgtctc	900
ttcagatatt	taacacaact	ttcaggcctt	cttttgcatt	cttttcaggc	cacaggacac	960
tggttttcgg	tgttcgttcc	cccaaccccc	ccaaaccacg	tatttttctc	atttggctta	1020
ttgcagtgc	tctctaaagg	tctgcctgct	ctgttcccta	ttcttcaccc	agcagccaag	1080
gccacgctta	aaatacaagt	caaatcactt	cattcctctg	ctcagaacac	ttctgcagtt	1140
tttctgtgca	cttgaggtaa	aaccttaagt	caacttgaag	attacagaat	tgaattgtct	1200
tactgaaatt	acctttccaa	aatcgatggg	tcccaactcg	tgccgaatcc	gatatcaagc	1260
ttatcgatac	cgtcgac					1277

<210> 46

<211> 442

<212> DNA

<213> Homo sapiens

<400> 46

gaattcggca	cgagggtga	gtgggggcca	tcctttttcca	tggttgagt	agggccatcc	60
tttcatgtgg	ctgagaggga	tccatccttt	cctgtggctg	agcgggatcc	attctttccc	120
gtggctgagt	ggaggcccat	ccttaggtac	ctccagtga	cagcctacat	ccttgaggga	180
tggaaaattg	atttgctct	tcacagactt	ctctggttct	agctttgggc	tatttatgctg	240
tgaagctgct	aagaacattt	cccaaatgtg	atttgtataa	atacctagga	gtaggattcc	300
tagtgtagg	tcagagtctg	tctagcagca	taaaaacaac	caaaagattt	tctgagcctc	360
agggcacatg	gggtggggccc	agagagcacc	caggaccgaa	gctgcccagg	acctccctcg	420

agggggggccc gtaacccatt cg

442

<210> 47

<211> 890

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (818)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (819)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (829)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (859)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (887)

<223> n equals a,t,g, or c

<400> 47

cttgtaaatg	tttcttttcc	cttaaataca	gataattcat	ttgtattgct	tattttatta	60
tgagctacaa	caaaaggact	tcaggaacaa	gtaatgtatt	agtatgggtc	aagattgttg	120
ataggaactg	tctcaaaaagg	atgggtggta	ttttaaatat	aaatagctaa	tggggggtgg	180
aggcctataa	aattaaatgc	cttgataaaa	atccaaaatg	aatgcaaaat	tgttttcact	240
tgtattgact	ttatgttgta	tgattccaat	ctctgttctg	tttggcactt	gtatttaatt	300
cttcaccttt	gtaagacatt	tgtatattgt	ggatgtgttc	attcaagcta	tttaatatct	360
ggcactgtta	atacacagta	ctttattgta	cagactgttt	tactgtttta	attgtagttc	420
tgtgtacttt	ttttggatgg	ggctggcatg	ttttctttgt	ttcctggcaa	tacgacgtgg	480
gaatttcaat	gcgttttgtt	gtagatgcta	acgtgtcaga	atcctttaca	ttcaactttt	540
ctaagaaaag	catttttcagt	cttgtagtgt	gtgcttacag	taactaattt	tgttgaaaat	600
ggtttcaagt	tattcaaatt	tgtacaggac	tgtaaagatt	tgttgacagc	aaaatgttga	660
agaaaaaagc	ttatagaata	aaagctataa	agtatatatt	aggatctgca	aacaatgaag	720
aattatgtaa	tatattgtac	aaatgtaagc	aaaggctctg	aaataaaatg	ccatagtttg	780
tgaaaaaaa	aaaaaaaaa	actcgagggg	gggcccgnna	cccaatcgnc	caaaagtggag	840
tcgtattaca	attcactgng	ccgtcgttta	caacgtcgtg	actgggnaaa		890

<210> 48

<211> 737

<212> DNA

<213> Homo sapiens

<220>

<221> SITE
 <222> (736)
 <223> n equals a,t,g, or c

<400> 48

gctgcgagaa gacgacagaa gggggagctc accagcgcca ccgtccccgg cgaagttctg	60
cgctggtcgg cggagtagca agtggccatg gggagcctca gcggtctgcg cctggcagca	120
ggaagctgtt ttaggttatg tgaaagagat gtttcctcat ctctaaggct taccagaagc	180
tctgatttga agagaataaa tggattttgc acaaaaccac aggaaagtcc cggagctcca	240
tcccgcactt acaacagagt gcctttacac aaacctacgg attggcagaa aaagatcctc	300
atatggtcag gtcgcttcaa aaaggaagat gaaatcccag agactgtctc gttggagatg	360
cttgatgctg caaagaacaa gatgcgagtg aagatcagct atctaagat tgccctgacg	420
gtggtaggat gcatcttcat ggttattgag ggcaagaagg ctgcccagg acacgagact	480
ttaacaagct tgaacttaga aaagaaagct cgtctgaaag aggaagcagc tatgaaggcc	540
aaaacagagt agcagaggta tccgtgttgg ctggattttg aaaatccagg aattatgtta	600
taacgtgcct gtattaaaaa ggatgtggtg tgaggatcca ttccataaag tatgatttgc	660
ccaaacctgt accatttccg tatttctgct gtagaagtag aaataaattt tcttaataaa	720
aaaaaaaaa aaaaanc	737

<210> 49
 <211> 571
 <212> DNA
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (249)
 <223> n equals a,t,g, or c

<220>
 <221> SITE
 <222> (548)
 <223> n equals a,t,g, or c

<400> 49

tgcaccacg cgtccggggg cgtacgcggg caagatggag gcgactacgg ctgggtgtggg	60
ccggctagag gaagaggcgt tgcggcgaaa ggaacggctg aaggccctac ggagagaaac	120
cgggcgcaag gtgagaagtg tggagttagg gtcgcagttg aggcgtccag cgttcgggggt	180
ccgggtcgcg cttgaggaga gcaaagggtt aataaggaaa gacagctgcc gagggcgcg	240
atgccgggnc gctaacgcat gcgcgagaag acgggcgccc tcccacgatg tctggggctg	300
cttggcggtg gactcctctg gcgctggtgc ggtcgtcgcg cacgcgcggg ggtgggcaar	360
gcatggtcag cgaccgcag tccatctgac tccgtcttcc cgggtgttgc tctgttaggt	420
atctagggtt gcctgtaggt tcagatgctt gttggggttag gcgtgatttg ttccgttcc	480
ctatggccta gctgggtctt aacccccgcc ttcgattctg agtcagacag actccccagt	540
tcgggcangc aattcccttg gaacaagggc a	571

<210> 50
 <211> 356
 <212> DNA
 <213> Homo sapiens

<400> 50

ccacgcgtcc gtaaaactcc acatttgtct gcatcagggg aaatgcatgg gcacacatcc	60
tccctccctc cctctctact ctctccctt ccttcaggcc tcttagcatt gtttgtttc	120
ccatttctga tactactact ccatgctgaa gatttgccat attactattt tggaacatt	180

gagtgataga	actcctagaa	aatttgcaaa	gaaatgttac	atactgtata	tcaaactctc	240
agattctagt	gttgaaaaag	tagcctatac	tttgctatta	cttataacctg	ctgccataga	300
aaaaaataa	gtttattcat	gacacattta	catttgatca	taaaaaaaaa	aaaaaa	356

<210> 51
 <211> 913
 <212> DNA
 <213> Homo sapiens

<400> 51	
ggcagcagtt	agtgtcctga aacgcctatg acagtgcctg gccaggggtt gggcacttac 60
tcatgttatt	cacttattca tatgtgggaa tgaacagttc agccagtata aaggacaggt 120
gagtacacac	caaggagatc ttcagggaga gatcagctgg ttttctctgt aagatgtcag 180
cttcatccct	ccaccgtctc cctgtgctca tggctctgtt ccctttccag gctgctgctg 240
ctggttcttt	gggtctgcag ccacctcca ccccatgaa gggcaaacc cgcattatgt 300
tacctctca	gtataaaaag agagagggtc tcaaaaaaaaa atccaaaaag 360
ttgctcttgt	cagctttggg agggcagact ccatagtgg agatgggctt ccaaccaacc 420
aaggagataa	atgccagagg gagcgaacca tgcagggtc aaagcacatc tctcccaaa 480
ctccccaggt	ggggaagcag gccagaggct ccacaaacc ctcaggagg cctggagtcc 540
agatgctgta	ctccagtac taaacaatca ctgaatctta aagctgacag gttcaaagct 600
cttactttgg	gccgagcgca gtggcttac cctgtaatcc aggcactttc ggagctgagg 660
tgggtggatc	acctgagggtc aggagtttga gaccaacctc gccaacatgg tgaaaaccca 720
tctctactaa	aaatacaaaa attagctggg cgtgttgaca cgtgcctgta atcccagcta 780
ctcggtaggc	tgaggcagaa gaatcgcttg aaccaggag gcagagggtg cagtgaagctg 840
agatcatgcc	actgcactcc agcgtgggtg acagagtgg actcccgtct tgggaaaaaa 900
aaaaaaaaaa	aaa 913

<210> 52
 <211> 1356
 <212> DNA
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (1231)
 <223> n equals a,t,g, or c

<400> 52	
cccacgcgtc	cgaagaatg ttgtggctgc tcttttttct ggtgactgcc attcatgctg 60
aactctgtca	accagggtgca gaaaatgctt ttaaagttag acttagtata agaacagctc 120
tgggagataa	agcatatgcc tgggatacca atgaagaata cctcttcaaa gcgatggtag 180
ctttctccat	gagaaaaagt cccaacagag aagcaacaga aatttcccat gtcctacttt 240
gcaatgtaac	ccagagggtat cattctgggt ttgtggttaca gaccttcaa aaatcacac 300
ccttctctgt	gttgagggtc aatcagccat aagaatgaac aagaaccgga tcaacaatgc 360
cttctttcta	aatgacaaa ctctggaatt tttaaaaatc cttccacac ttgcaccacc 420
catggaccca	tctgtgcccc tctggattat tatatttgggt gtgatatttt gcatcatcat 480
agttgcaatt	gcactactga ttttatcagg gatctggcaa cgtagaagaa agaacaaaga 540
accatctgaa	gtggatgacg ctgaagataa gtgtgaaaac atgatacaaa ttgaaaatgg 600
catccctctc	gatccctctg acatgaaggg agggcatatt aatgatgcct tcatgacaga 660
ggatgagagg	ctcaccctct tctgaagggtc tgtgttctg cttcctcaag aaattaaaca 720
tttgtttctg	tgtgactgct gagcatcctg aaataccaag agcagatcat atattttggt 780
tcaccattct	tcttttgtaa taaattttga atgtgcttga aagtgaagaa caatcaatta 840
taccaccaa	caccactgaa atcataagct attcacgact caaaatattc taaaatattt 900
ttctgacagt	atagtgtata aatgtggtca tgtggtattt gtagttattg atttaagcat 960
ttttagaat	aagatcaggc atatgtatat attttcacac ttcaaagacc taaggaaaaa 1020

taaattttcc	agtggaggat	acatataata	tggtgtagaa	atcattgaaa	atggatcctt	1080
tttgacgatac	acttatatca	ctctgtatat	gactaagtaa	acaaaagtga	gaagtaatta	1140
ttgtaaattgg	atggataaaa	ttggaattac	tcatatacag	ggtgggattt	tatcctgtta	1200
tcacaccaac	agttgattat	atattttctg	natatcagcc	cctaatagga	caattctatt	1260
tggtgaccat	ttctacaatt	tgtaaaagtc	caatctgtgc	taacttaata	aagtaataat	1320
catccaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaa			1356

<210> 53
 <211> 1547
 <212> DNA
 <213> Homo sapiens

<400> 53

ggcacgagcg	gctgcgggcg	cgaggtgagg	ggcgcgaggt	tcccagcagg	atgccccggc	60
tctgcaggaa	gctgaagtga	gaggccccga	gagggccccag	cccgccccgg	gcaggatgac	120
caaggccccg	ctgttccggc	tggtgctggg	gctgggggtcg	gtgttcatga	tcctgtctgat	180
catcgtgtac	tggtgacagcg	cagcgccgcg	cacttctact	tgacacagtc	cttctctagg	240
ccgcacacgg	ggccgcccgt	gcccacgccc	gggcccggaca	ggacagggag	ctcacggccg	300
actccgatgt	cgacgagttt	ctggacaatt	tctcatgctg	gcgtgaagca	gagtgcctt	360
cccagaaagg	agacggagca	gccgcctgcg	ccggggggagc	atggaggaga	gcgtgagagg	420
ctacgactgg	tccccgcgcg	acgcccggcg	cagcccagac	cagggccggc	agcaggcgga	480
gcggaggaac	gtgctgcggg	gcttctgcgc	caactccagc	ctggccttcc	ccaccaagga	540
gcgcgcattc	gacgacatcc	ccaactcgga	gctgagccac	ctgatcgtgg	acgaccggca	600
cggggccatc	tactgctacg	tgcccaaggt	ggcctgcacc	aactggaagc	gcgtgatgat	660
cgtgctgagc	ggaagcctgc	tgaccgcggg	tgccgccctac	cgcgaccgcg	tgcgcatccc	720
gcgcgagcac	gtgcacaacg	ccagcgcgca	cctgaccttc	aacaagttct	ggcgccgcta	780
cgggaagctc	tcccgcacc	tcatgaaggt	caagctcaag	aagtacacca	agttcctctt	840
cgtgcgcgac	cccttctgtc	gcctgatctc	cgccttccgc	agcaagttcg	agctggagaa	900
cgaggagttc	taccgcaagt	tgcgcgtgcc	catgctgcgg	gtgtacgcca	accacaccag	960
cctgcccggc	tgggcgcgcg	aggccttccg	cgtggtgctc	aaggtgtcct	tcgccaactt	1020
catccagtac	ctgctggacc	cgcacacgga	gaagctggcg	cccttcaacg	agcactggcg	1080
gcaggtgtac	cgcctctgcc	acccgtgcca	gatcgactac	gattcctggg	gaagctggag	1140
actctggacg	aggacgccgc	gcagctgctg	cagctactcc	aggtggaccg	gcagtcgcgt	1200
tccccccgag	ctaccggaac	aggaccgcca	gcagctggga	ggaggactgg	ttcgccaaga	1260
tccccctggc	ctggaggcag	cagctgtata	aactctacga	ggccgacttt	gttctcttcg	1320
gctaccccaa	gcccgaatac	ctcctccgag	actgaaagct	ttcgctgtgc	tttttctcgc	1380
gtgcctggaa	cctgacgcac	gcgcactcca	gtttttttat	gacctacgat	tttgcaatct	1440
gggcttcttg	ttcactccac	tgctctatc	cattgagtag	tgtagcgata	ttgtttttta	1500
agattaatat	atttcaggta	tttaatacga	aaaaaaaaaa	aaaaaa		1547

<210> 54
 <211> 1338
 <212> DNA
 <213> Homo sapiens

<400> 54

cccacgcgtc	cggttcccc	atctgtctct	caggagcgag	atctgatcgc	tgaatttgcc	60
caagtcacaa	attgggtccag	ctgctgcttg	cgtgtctttg	catggcacc	ccacaccaac	120
aagtttgcag	tggccctgct	agatgactca	gtccgtgtgt	ataatgccag	cagcaccata	180
gtccctctcc	tgaagcaccg	gctgcagcga	aatgtggcgt	ctctggcctg	gaagcccctt	240
agtgcctctg	tcttggtctg	ggcctgccag	agctgcattc	ttatctggac	cctggaccct	300
acctccttgt	ctaccgacc	ctcttctggc	tggtgccaa	tgctgtctca	ccctgggcat	360
acacctgtta	ccagcttggc	ctggggcccc	agtggggggc	ggctgtcttc	agcttcaccg	420
tggatgctgc	tatccgggta	tgggatgtct	caacagagac	ctgtgtcccc	cttccctggg	480
ttcgaggagg	tggggtgacc	aactgctctg	gtccccagac	ggcagcaaaa	tcctggctac	540

cactccttca	gctgtctttc	gagtctggga	ggcccagatg	tggacttgtg	agaggtggcc	600
tactctatca	gggcgctgtc	agactggctg	ctggagccca	gatggcagcc	gactgctgtt	660
cactgtattg	ggagagccac	tgatttactc	cctgtctttt	ccagaacgtt	gtggtgaggg	720
aaaggggtgc	gttgagggtg	caaagtcagc	aacgattgtg	gcagatctgt	ctgagacaac	780
aatacagaca	ccagatgggtg	aggagaggct	tgggggagag	gctcactcca	tggctctggga	840
ccccagtggtg	gaacgtctgg	ctgtgcttat	gaaaggaaag	ccaagggtac	aggatggtaa	900
accagtcctc	ctcctttttc	gcactcgaaa	cagccctgtg	tttgagctcc	ttccctgtgg	960
cattatccag	ggggagccag	gagcccagcc	ccagctcctc	actttccatc	ttccttcaac	1020
aaagggggcc	tgctcagtg	gggctggtcc	acaggccgaa	ttgccccat	cccgtgttac	1080
tttgtcaatg	ccagtttcc	acgttttagc	ccagtgtctg	ggcggggcca	ggaacccct	1140
gctgggggtg	gagggctctat	tcatacctg	cccctcttta	ctgagacatc	cccaacctct	1200
gccccttggtg	accctctccc	agggccacca	cctgttctgc	cccactcccc	acattcccac	1260
ctctaagaat	aaataagttt	tccttttgtt	ttccaaaaaa	aaaaaaaaaa	aaaaaaaaaa	1320
aaaaaaaaaa	aaaaaaaaaa					1338

<210> 55

<211> 2071

<212> DNA

<213> Homo sapiens

<400> 55

cgccacgagc	caaaacaaaa	atthtgtgca	gtcctgtcat	ttaatcatct	tatcagtggc	60
ccaagggtgac	ttttttcaag	agctctgttg	ataccatgtg	tgtgtgttat	tttctggtgt	120
ttttacagat	ttggggccagg	cttagtcatc	tattggtatg	gatttatcca	ggagctggac	180
tgcaaccggg	aaaggggcat	cctgtctaaa	gcctgtttcc	ccacgaacat	tgtaacctta	240
tgccacagca	tagcttgacc	ctgaagatcc	tggagagaaa	gctggggagga	aaaggtgaat	300
ccggaagcaa	ttttactttc	ctgcaactgt	agatcctggc	aacatctgcc	ctgaacttca	360
gctgaactct	tgctgcccgt	agtcacacca	ctacctcttt	agacaaacat	atcaagagtt	420
tctgttttcc	ttcatccctt	gctgatgtga	acagccaaga	actacagaca	caaccacttc	480
attatcagca	tttctgtctc	tgtaaaacaa	ttagtttata	gatagtcata	atctttcttt	540
cttcgggatg	gtttatccct	gtatgctgaa	caaaagaaaa	aatgtgaaga	ctgaaagggtg	600
tgatttttca	attctcctcc	cagttaccga	atcaccctcc	ttattttttt	ttccctaaagc	660
ctccgttcac	tgctctctcc	ctctcctttc	ttttccatg	ttgcaactcc	cagagaaccc	720
agcattgcat	taccatogtg	gaataatcta	gcgcaaacct	aggaaagctg	aagccacaaa	780
gtccaaagcc	acctttgtac	tcacctgcag	agctccagaa	gaccttgatg	gcagcctgcc	840
tatgctgtgt	gtttgtctata	ttcaatcttt	acggcttcc	gacttctgtg	acagtaagcc	900
aagtgcacaaa	atacacttga	tgagaatttc	ctcttttaat	aatgttattt	gaacaccaca	960
tatttttagat	ttatcttatt	tgaaagtatt	agttccattg	tgcttggaag	ccacactcct	1020
ttagattggg	ggccgagagg	cgacaacca	acattgagga	gagtttattt	ttaaacatgg	1080
ctagtgtgca	gtatgtacgt	gagctagtat	ttttatgagt	cgagtttttt	aaaggcacat	1140
tctgtatact	gcttagtata	tgcatthtat	accatgtaat	tataaaacac	tcgagtaagt	1200
tcagcattag	aatgttttag	ctttgtatga	actgagtggt	ccagaaataa	acctggagca	1260
attttttaat	aagcaaaata	aaggagattt	ttctatttgt	tcactttaat	ttattcactt	1320
ttgtgtactt	ttatgtactg	caaatcagat	ttcagtctaa	agcgaaacat	caataagtta	1380
ataataaacc	tatcttttgg	gaagtgtgaat	attaatctgt	acccaaaacg	tatttagtaa	1440
aatatttgcc	cccgccaccc	tgccatgctg	acataacaac	ttttataatg	ttgaatagat	1500
gatattgggaa	atactaataa	caacaatgta	atttttgcag	acagctttta	cttatataca	1560
ttggttgatt	tttttcaaaa	gactaaatat	gtcattttata	ctttgtttat	tttctaccaa	1620
agaagggttg	taaaaatatg	cctgtgctt	ttccttttga	aggacacaaa	cctggtccca	1680
acatgtgtgg	atttttaactc	tgagtgggtg	gcattaaatc	aaaagagaga	ggcagaagat	1740
gaaatgctaa	agaagggtca	ggcaaacctc	tgtttcagta	taaaattcat	catgcaggct	1800
tctgagtgaa	atagaatgat	ttgaaccac	tactgtattg	cctggatata	cacacacaca	1860
cacacacaca	cacactttat	acaaaaatgt	taaaagcagg	tttctgggca	tgttctaaac	1920
tgttttttct	ttaggaataa	attacattta	tctgcacaga	tgttgaaaat	cctgttaaac	1980
ccttgtaag	gatttgttta	ttttacatta	aacaaattta	ttatgatgaa	cgtgaacaaa	2040
taaattaaaa	aaaaaaaaaa	aaaaaaaaaa	a			2071

<210> 56
 <211> 1899
 <212> DNA
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (1439)
 <223> n equals a,t,g, or c

<400> 56
 ccacgcgtcc gccacgcgt ccgcccacgc gtccgttacg atttataaaa gcaaactttt 60
 aattttttcat aatatatcta ttccctaccta ggttttttta atcaaattaa taaaatggta 120
 ctccctttttg tgttattgtt cagaccaaatt tttatcagtg tccttcaccc tttattctac 180
 tcacattgtt tattttctata cttaataagt cctgttcact ctccctctat aatatattac 240
 aaacctgacg attgtcacta caccctattc attcctgggc tactacaata gccagggtcta 300
 cctactttcc ttctggcatt tggacaatct gtgtttctgg aagtagccca aatgatcttt 360
 ttttgttctt atattattca gatcctatta ctcttgatc tcaaccctca caagatattc 420
 ctattatatg cttaataaaa tccaatatcc ctacctggc ctgtatactc aataaagtcc 480
 aatatcccta ccttggccta caagatctta gatgactggg atctaggaaa cctattcagc 540
 taatttcctg caattctcta tgtttacctg ctcaagggtt ttaccactcc tcaaataatc 600
 aaactgggtt tgcattcttt tagttctatc tattcttttg cagatgccaa tcttagtaat 660
 cataagcagt ttacctctaa attaatgcta aatattgtaa aagggtgtatt ataagttcag 720
 tctttcattt attctttcaa ccaccaagtg tcgattgaac acatacattg ggtcacacca 780
 atcaaataag atagactccc aggcattccc gagcatctac tactcaatc atctcaggat 840
 ctatacaagg ggtcattttta tcagaaacca aagtcttgat gctgtccgaa aaatcacggt 900
 ttcaccatga tctcctactt ccctaggtgg aagtataact ttttaggata tatcatgggt 960
 gctgacttaa cttttgtatt ttttaaatat actcatgaca agtatcatat aaaacctaac 1020
 cagcaacttt gcaccagcaa aagtttttca acatttcaat tcttacaata tcaaatgata 1080
 taatttctta tgtagtaaaa aattcacaca tctgcaaagc ttggtttcac taccacctgt 1140
 taaaatctta cttttggaag ctatttatga ttgaaaaaca ctttacctca ctcaaaaaga 1200
 gctggaagtc tctcttcaat ccaatatgca cacagaagac aaaaagctgt atcatctctt 1260
 gatgatatat ttgaaatcat atggccacgt ctgtccattg tcttcagagt ttctaagtat 1320
 ttcagaaaaat tatgacttgc actgtagaac tattttaaag aaattccatg gtgcaaacag 1380
 aaaaactaaa acttttcatg ttaggataat ttattaaaaa tacaacaaaa tcttatgtnt 1440
 acataaaga atagtaacta gcctttttga gagggaaatt tttctctcat aacttctttt 1500
 ctagtaattt caataaagaa taactgccat tccaacgttt agcccatctc actctcttgt 1560
 cttcttatgg ccaagtattc aagcttgaaa tttgcagagg aaattcttgt ccgtttttta 1620
 tatcatgtgg taagcctaatt aaaacatctt ctgaaataat tagcccttaa aaggatagta 1680
 tcttctacct gacagaggca aatattattg aaaagtttgt acctataag cacattaatc 1740
 atggagtcct ggaactggat tctgtctaag actgactttt gcttaattaa gtccacagag 1800
 attttccaca tatttttcca gaacattgca tgtagagata ttgtcgatca atcacataac 1860
 tagggtcaga aagatgtaac aaggagagaa aaaaaaaaaa 1899

<210> 57
 <211> 1543
 <212> DNA
 <213> Homo sapiens

<400> 57
 ggcacgagat ttgattctca tgctcctttc aaaagagcat actagtttgg ggtggttgggt 60
 tattttctta accttagcaa gccagcttat ttccctatgga agcagaactg gaaacagcag 120
 atgtccacca tgcttataca ggacactaca cactgtctcg acaagccatg tcttttcttc 180
 cctcttcgtg agcactttct ctggtgatga gttagtatgg actacttgaa cctcaaaact 240

gggcctctca	cccaaagcca	aatgaagtag	cgtatgccag	gatgatgttt	cttttggggc	300
gttggcagtg	agactgctaa	gcaggctgcc	ttaggttttg	ctgtggcaat	gctagcagat	360
tgttccctct	ttcaaagggg	caaaaatata	attttgggtat	gataactgac	tttctattta	420
cagtttcttg	cccccaaaga	caaaccaagt	ggagacacag	cagctgtatt	tgaagaaggt	480
ggatgatgtg	acgatttagt	aagtactttt	aatatgcacc	tggtgttctg	tgattgaagt	540
cacctgagct	gtaaatacag	ccacaaaggc	tgattatctt	acacttggtg	cttatttgtg	600
ttttaatttc	caatacacca	gaagcttcct	acaccattat	atattgccat	tataaattca	660
atcagatagg	taatttcata	atagaaattc	ctgtgtttca	tggtgtcggc	tatattgttc	720
attcagatta	atcctctccc	ttgaagggct	gaaaaagact	aggagagctat	tccattagta	780
gcaaaatggt	gtaattcact	gaaattgctg	ttaaccaaaa	ataagtaata	caacatggca	840
ttttgtgttg	gttgacaaat	gaaacaggcc	ttaaaagggc	tacttcttaa	atgttctcaa	900
ttaacttaat	gtaaacaaaa	tagaccgata	ggcatttgag	gatttctgga	ccccattaca	960
ccatgttgtt	gatgtctggg	aagctgtgta	gtaaatgtct	ttgttatcta	tccttaattgt	1020
ttggaaactt	cccgcttta	agcttcata	gacaactgac	caacaaacac	tacgtactat	1080
gatgtcaatc	tttttttagag	acattctcat	tactaaaatg	agtggatact	tgaatgttta	1140
actcctaaaa	taatgagcgg	tgaataaatg	agcaagtaca	tgcatgcctt	ccaatgtaga	1200
gtcattttca	ttaaaccctc	tctcaccaga	gaagcagtg	tatgaaattg	gcctgattcc	1260
tttctaagtg	tggtgttctt	gttcacagtt	ggacatgata	taggtcgtgg	atgtatgggg	1320
aatctaagag	agctgccatc	gctgtgatgc	tgggagttct	aacaaaacaa	gttggatgag	1380
gccattcaag	gggagccaaa	atctcaagaa	attcccagca	ggttacctgg	aggcggatca	1440
tctaattctc	tgtggaatga	atacacacat	atatattaca	agggataatt	tagaccccat	1500
acaagtttat	aaagagtcac	tgtaaaaaaa	aaaaaaaaaa	aaa		1543

<210> 58

<211> 1133

<212> DNA

<213> Homo sapiens

<400> 58

ggcacgagct	ggagcaaaga	tattgtttga	aggagagttt	atggtttttg	attttaaacg	60
ggcaggggtc	tttttcctct	cattttttgtg	gacaagagag	gccttcgcct	ttattttttac	120
tctccctctt	ctgctgtccc	tgtgcagagg	aaaaatgaag	aattctccca	gaagtgactt	180
gtcaagactt	aaaaaaaaatg	tttttaatgc	atttcttctt	tgtctagtgc	ctcggtttat	240
ctctaacagg	ggctgtccag	tatatcggtc	ctgttaggag	gggagaaaaa	gttcttccaa	300
aggctggaga	agtgaacaag	gagtc aaatt	tattttccca	attcaacttc	ataattatca	360
tttctttggc	ttcatgtctt	cccgtaactc	atgtgggttg	gatccatccc	atctgggtca	420
cttcagtcta	cttcacgtac	ttgaaaaggc	tttccctttac	acttccagga	ccaaacagca	480
acttcctgcc	acacacttcc	accctatcac	tgggagaaat	ccttttctgg	acatgagcct	540
ttgacctggg	tggggcagaa	agaaccacaa	actccatctc	ccaatagaac	tttgaaattc	600
actcagcttt	tcctttcatg	ctgtttgttg	cctgcttggt	gcactcctcc	tgccccagaa	660
ctgcaagatt	tttagcttca	cccctttctg	agagtaattg	tatcttttat	cagaatcagt	720
atcagttccc	ctgtattctg	tgcttcacgc	aatttgcaag	actgacctct	tttaagcatt	780
taattcactc	ccagagtcac	ctggtcaggt	tgcaatatga	ggacttctct	gtctcctctg	840
aagcctggga	caactgagctt	acttaataca	ttagatgttc	aaaagaggag	cggtgtttca	900
tctttcaaaa	tgtaggcca	ttactttgag	tataaaatcg	acttattaat	gattagtaat	960
ttttctaaag	tattgggaaa	actttcttat	tttataagat	cttaacaagc	ttaaaaaaga	1020
attttatgac	cagaatccaa	caagagctct	attttggaat	tgtgccaag	ttggtgatgt	1080
ttactctaaa	attaataata	aaactacttg	taagcaaaaa	aaaaaaaaaa	aaa	1133

<210> 59

<211> 1490

<212> DNA

<213> Homo sapiens

<400> 59

ggcagaagtt	cctctgcg	tccgacggcg	acatggg	ccccacggcc	ccggaggccg	60
gcagctggcg	ctggggatcc	ctgctcttcg	ctctcttcct	ggctgcgtcc	ctagacatca	120
cggctgcagc	cctggctacg	ggtgcctgca	tcgtagaate	ctctgcctcc	ccctcatcct	180
gctcctggtc	tacaagcaaa	ggcaggcagc	ctccaacgcg	cgtgcccagg	agctgggtgcg	240
gatggacagc	aacattcaag	ggattgaaaa	ccccggcttt	gaagcctcac	cacctgcca	300
ggggataccc	gaggccaaag	tcaggcacc	cctgtcctat	gtggcccagc	ggcagccttc	360
tgagtctggg	cggcatctgc	tttcggagcc	cagcaccccc	ctgtctcttc	caggcccccg	420
agacgtcttc	ttcccattcc	tggaccctgt	ccctgactct	ccaaactttg	aggatcatcta	480
kccwkctgg	gggacagtgg	gctgttgg	ctgggtctgg	ggcaggtgca	tttgagccag	540
ggctggctct	gtgagtggcc	tccttggcct	cggccctggg	tcctccctc	ctgctctggg	600
ctcagatact	gtgacatccc	agaagcccag	cccctcaacc	cctctggatg	ctacatgggg	660
atgctggacg	gctcagcccc	tgttccaagg	attttggggg	gctgagattc	ttccctagag	720
acctgaaatt	caccagctac	agatgccaaa	tgacttacat	cttaagaaat	ctcagaacgt	780
ccagcccttc	agcagctctc	gttctgagac	atgagccttg	ggatgtggca	gcacagtggt	840
gacaagatgg	acactggggc	accctcccag	gcaccagaca	cagggcacgg	tgagagact	900
ttctccccgt	ggcagcact	tggctcccc	gttttgcccc	aggctgctct	tctgtcagac	960
ttctcttttg	taccacagt	gctctggggc	caggcctggc	tgcccactgg	ccatcgccac	1020
cttccccagc	tgcctcctac	cagcagtttc	tctgaagatc	tgtcaacagg	ttaagtcaat	1080
ctggggcttc	cactgcctgc	attccagtcc	ccagagcttg	gtgggtccga	aacgggaagt	1140
acatatggg	gcattggtgg	ctccgtgagc	aaatgggtgc	ttgggcaatc	tgaggccagg	1200
acagatgttg	ccccaccac	tggagatggt	gctgagggag	gtgggtgggg	ccttctggga	1260
agggtgagtgg	agaggggcac	ctgcccccg	ccctcccat	ccctactcc	cactgctcag	1320
cgcgggccat	tgcaagggtg	ccacacaatg	ttttgtccac	cctgggacac	ttctgagtat	1380
gaagcgggat	gctattaaaa	actacatggg	gaaacaggtg	caaaccctga	aaaaaaaaa	1440
aaaaaaaaa	aaaaaaaaa	aaaaaaaaa	aaaaaaaaa	aaaaaaaaa	aaaaaaaaa	1490

<210> 60

<211> 1336

<212> DNA

<213> Homo sapiens

<400> 60

ggcagcaggt	ccagccaagt	tatctaccct	caaatctctg	actaatcatg	tttgtgtttt	60
gggtattttaa	aattacatac	atatatatcc	tttttgccaa	aaacaaaagt	cttgcttctt	120
gtcaaatgat	tgctaaagta	gatctttacat	tttttgttat	tatgtatata	tttatacaca	180
cccccaacac	acttagtgat	ttctgttatt	tcctagggag	cacagcttta	aggctatgag	240
atacaactaa	aaggagccca	tctatttggg	ttccagcca	attattgtac	tcacatttca	300
ggggagaatc	tgaaattcct	gtcatgttta	cagcaacaat	ctatcattcc	tggttagctc	360
tcagcctctc	tctccttoca	taggttagaa	ttatgtcatt	ttgttactta	gtggccacgt	420
ctattttctga	gaaagactgg	ttacattttat	gtggcatctc	aggatcatt	aaggaaaagc	480
cagagcaggg	gtgagcagag	gtcaaaaacca	cagacgcagc	agggccattt	gccgcctttg	540
gccgggatca	caaccactgc	agtctcccag	caggtagggc	ttgccaaagcc	taaggctccc	600
catccaatct	agacagaggg	gcgctcagag	cagactttgc	cgtagcccat	gtctgggtgag	660
cacaacaggg	aatgaattgg	gcactccact	ccccgtctc	tctggcccag	ccctgaacta	720
gatgagctgc	atttcatgga	gcccatttta	aaacctcttt	ccttatgact	ttgttactca	780
agtccagagt	tctctgtgca	cttctgctag	ataaggagt	taagccctgc	ccccagcac	840
tggcagcacg	ctgggcccctc	cccacacagg	acaccgtgca	gttcggggg	aagctgactc	900
aaatcaacct	tgaaatctca	tgaaaacaaa	atgacttgct	tttttatttg	atagtgtaat	960
atcattttat	aaatttttta	gggtttttct	cggtgttaata	ttgtacagtt	ttgcatggcc	1020
tggtgtgata	attttttggg	tagaatataa	tgctgacaaa	tgtggatgga	ggggaagata	1080
ctgcttttagc	ctatcactcc	ttattttatt	ttgtttgggt	ttatgccctc	agtgctcttag	1140
ggaactttat	aagagatcct	ctgctaccaa	acaatgatgt	ggattctttt	gcacagaaat	1200
atttaagggtg	ggatggtaaa	aaatgtcaca	aaagactcct	caccaatact	ttatgttgat	1260
atcacttaat	attaaccaga	ctttgctgta	ttgcaataaa	acagagaact	gttaaaaaaa	1320
aaaaaaaaa	aaaaaa					1336

<210> 61
 <211> 1705
 <212> DNA
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (779)
 <223> n equals a,t,g, or c

<400> 61
 gaattcggca cagctttaca gtacatagga atttgagaac cacttcacag gaagagggaa 60
 acagcccaat atttatttat gtatacacat aatcccaagt gtgtgctggg gccaccaggc 120
 ttacctgggg gaacaaggac tgcgtgcat gtgagtgcag acattaatag catttacata 180
 ctgtacagat gcaacctttg atgatacata tatttgataa aaatgagaaa acagatttgt 240
 tgtagagtac ctgtccactt ttatagcatg agaacagtac aatcaactat ttattttgca 300
 gttactcatt tcagtgattg agaatttctg tgcgtgacag agagacggcc tgtaattggg 360
 ctcatcatcc acttgattct aacatgatct ctgcccagaag ttccatttct tgggctttga 420
 tatttataat ggcgcctgct ctctactctg tcttacgctt tccgagcaag ttcaaaccag 480
 aaagaaaagg tgaggctaga agcccaaagt gaggtagtgt gaggaccaca aggaagccca 540
 ccactccaca gtagatgatc aaaaccacat cctcacgtgg gaggtagcac ttgggagagg 600
 gtgtagtctg tgggcgtgat gctaccctgg aaaggaraag ggaaagtatt gctgagagca 660
 ccaggcacag gttgaacacc gcagtcttag aaacagcaga gggaagactg cyttctcagg 720
 tccccctcag gtgaggcagg gaacggggccc tctcacctg agaccaaggg ggcccagcnt 780
 tctccctgca cagctcacc cgcaccagcc caggctccag caggagagac aagtaaggcc 840
 caagtgtgcc tgagtggaaa atgtctggga cactgacctg tcaaaaactgg cccctggctc 900
 actgggttcc catcaaatat agtgggggat ccataacaga gttsagagag gcaccgtgga 960
 gttccagggt catcggtcag cgaggaacaa ggagggaag gtgtcttctt gccctttgat 1020
 gctcaamtaa gcatctgttc cctagaaata catgtgtcca ggtcgtctcc atgggctttt 1080
 ctttgcagat acttttatgt ggaacaacag tggcaaatgt ttctatttac ttgaatttg 1140
 aaaatggttag ggtttcctcc acctttttat gaagtaaaag amcctgtcgt accagcatca 1200
 tgagctggat gcaggagccc atggctgaaa ggagttaaaa cgcccagtgg tcattaagtg 1260
 aaacatcttt tatcaacctg caaaagctgc cagcgtttct tgccagggtca aatgggcatg 1320
 tttagaaaaat aagagaagat ggctgagtat agctaataa taaatgggtt ttcttttaga 1380
 aaattaaaca cacacagagt gtaagaggag aggatacggc cctccctgaa ggataaagtc 1440
 cmcctggacg gtgcccctgccc ctgcttcttc acattaaactg cccaggaatg tcatgctgat 1500
 tgggtcccggt aaggggtgtt ggcaaggggc agtgtatgga gctacgtgta gaaggagaga 1560
 aatttgtgtg tggcttttgt aaattttgac cgattgcagc aattaaatag ttgattactg 1620
 tgttgattta aatacttatg aaagctttca gacaaaaata aactttcacg ttaccatgaa 1680
 aaaaaaaaaa aaaaaaaaaa tcgag 1705

<210> 62
 <211> 1031
 <212> DNA
 <213> Homo sapiens

<400> 62
 ctgacagtca ccgtccggaa tcccgggtcg acccacgcgt ccggggcgtcc gcgtcggcga 60
 tccgggggtct gggcgcgggc aggtctgggtg tggcagatgt ggatagctgg accctcctgg 120
 gtgcctctgc gttatgtcgt ttggttgatg cattttggaga gaatctgtgc tctccacaac 180
 tgcaggggga acatgcttct ctggcctctt cagatcaggg tggctgttct tgggtgctgc 240
 accaaaactc cagcagtggg ctttcttcaa gtggccggt cactcattc ctgccaagac 300
 ccaggtccgt gttctcacag tgcctccatc tttcctccgt gtgagcgtgg gctgtgcgga 360
 gatggacctc ggtgtgtgct gggctgcgtg cactgccatc gctcccttct acacgagccg 420
 gcgtggaccc agggctgagc tgtgaccacg agggccatcc cgacgagccg ccatggaccc 480

agggctgagc	tgtgaccatg	agggctatcc	cgacgagctg	ccgtggaccc	agggctgarc	540
cgtagaccatg	agggscatcc	cgaaactgtg	attgttttct	gaatgtggaa	gcgtcggccg	600
agcgcgggtcc	ttggagaacc	cttgctgcgg	agaatgacgc	tactctgcg	gcctggcctc	660
gcctcmccat	cccttctgca	mactcacagg	acaggattga	tggggagagc	ccaggcgaaa	720
cgaagctgac	cctastgacc	agccctactg	gcggctctgg	aacaggcccg	ggaccgggtg	780
cctggagaca	gtggtccttt	ctggatgggc	tgagctctcc	gaaggcttcc	ccactggag	840
tcagcccttt	aggagtgtgt	ccagtgcctc	cagaaagctg	ttttttgggg	gacggaggac	900
ttggcctgga	attctggaat	tcccaggggg	tcagacatgg	ttatgggaag	tttaataaaa	960
ccggtkaatc	acgtgaattc	tgaaaaaaa	aaaaaaaaa	aaaaaaaaa	aaaaaaaaa	1020
agggcgccg	c					1031

<210> 63

<211> 1589

<212> DNA

<213> Homo sapiens

<400> 63

cgccacgagc	taagccataa	tagaaagaat	ggagaattat	tgattgaccg	tctttattct	60
gtgggtctctg	attctccaat	gggaatacca	agggatatta	tttttacaga	tggttttcca	120
tactggaacc	caaaggtaaa	gacactcaag	gacagacatt	tttggcagag	catagatgaa	180
aatggcaagt	tccctggctt	tccttctgct	caactttcat	gtctccctcc	tcttgggtcca	240
gctgctcact	ccttgctcag	ctcagttttc	tgtgcttggg	ccctctgggc	ccatcctggc	300
catggtgggt	gaagacgctg	atctgccctg	tcacctgttc	ccgaccatga	gtgcagagac	360
catggagctg	aagtgggtaa	gttccagcct	aaggcagggtg	gtgaacgtgt	atgcagatgg	420
aaaggaagtg	gaagacaggc	agagtgcacc	gtatcgaggg	agaacttcga	ttctgcggga	480
tggcatcact	gcaggggaagg	ctgctctccg	aatacacaac	gtcacagcct	ctgacagtgg	540
aaagtacttg	tgttatttcc	aagatgggtga	cttctatgaa	aaagccctgg	tggagctgaa	600
ggttgacgca	ctgggttcta	atcttcacgt	tggaaagtga	gggttatgag	gatggaggga	660
tccatctgga	gtgcagggtcc	accggctggg	accccccaacc	ccaaatacag	tggagcaacg	720
ccaagggaga	gaacatccca	gctgtggaag	cacctgtggt	tgagatgga	gtgggcctat	780
atgaagtagc	agcatctgtg	atcatgagag	gcggctccgg	ggagggtgta	tcctgcatca	840
tcagaaaattc	cctcctcgcc	ctggaaaaga	cagccagcat	ttccatcgca	gacctttctt	900
caggagcgcc	cagccctgga	tcgcagccct	ggcagggacc	ctgcctatct	tgctgtgct	960
tctgcggga	gccagttact	tcttgtggag	acaacagaag	gaaataactg	ctctgtccag	1020
tgagatagaa	agtgaagcaag	agatgaaaga	aatgggatata	gctgcaacag	agcgggaaat	1080
aagcctaaga	gagagcctcc	aggaggaact	caagagaaaa	aaaatccagt	acttgactcg	1140
tggagaggag	tcttcgtccg	ataccaataa	gtcagcctga	tgctctaata	gaaaaatggc	1200
ctcttcaagc	ctgggcctcc	cattggccaa	acacagcagc	aaaccagagg	acaagggagc	1260
ccagtggcac	tgtccataga	ggacagattc	ctgggggtcca	gaagaggggtg	gagaaagctg	1320
aaggctggag	agtgaatcta	gggcatataa	ggccccacac	agagcccagc	acagagacgg	1380
ccttgacagct	atcaggaaga	tgaggagctt	ccttcatggc	ctgctgtggg	ctgagtaaat	1440
aatatgattg	ccttctacag	cgctagagat	tcatatgttt	atccccattt	ttcagggtgag	1500
gaaatgcttc	agatgaggct	ccaccttggt	aaataaattg	gatgtatgga	aaaatagact	1560
gcaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa				1589

<210> 64

<211> 1088

<212> DNA

<213> Homo sapiens

<400> 64

ggcagagct	aagccaaccg	cactgaagga	gtggggagaa	gagcatacgc	caggagcctc	60
ctgectcaaa	gtgctcccct	aagtcttctt	cctcctgtgc	tgacctcagg	gtgggtctgac	120
ccttccctcg	gtgtggggga	tgtggccctc	tcaggtgccc	ctacttgctt	tctgcttctt	180
tctggtgaag	tccacctcca	acattaacct	gccccccca	cccccgctcat	ccctggagaa	240

ttccagcttt	gtcgtatctc	agagagggaa	tctaattgtt	tttggggggc	aaaagaaagc	300
aacgttttag	tatcacttct	acttggaccg	catgcctttt	tatagccaaa	tttctgtgta	360
tttcgtaaat	ggatttcgcg	ttaatggata	tttatgtaat	aactagactt	ctcagattat	420
tgtgagaagg	gtcagggttg	aaggggtgta	ggaagagggg	tgaggggtag	tttttttctg	480
ttctagtttt	tttttttttt	ttgtcatctc	tgagggtggac	cttgtcacct	gtggttattg	540
gggccaaggc	ggactcagct	cccggggaga	agggcctctc	tgccatttcg	gtcccaaggt	600
gagctgacac	aggcggttct	tttgggactg	tggaagcatc	agatgccagc	actgactcag	660
gaacagcaag	tcagggcaga	gaggaggagg	gaggctgtca	ggatggaaat	acctggactt	720
ttctttgctt	ccctcgcaaa	ctgggggtctt	ctctaccgaa	cttcccagga	tttcatctca	780
ccatatctgt	gtgcgcgccc	cagcaccccc	caccacacct	tgggggggccc	gtgagcgtgt	840
gtcttcattg	cctctctccc	cttggcgtct	gatgaccaca	gcaaagcact	gggaatttct	900
actcttcatg	cctcatcctg	cagcctcggg	ttcgcatctt	ctctttcttt	tcctctttcc	960
ctctttccct	gggattgact	ctgagtggaa	taccttggca	catccactag	gatctactgt	1020
ctgcactgtt	ttctttgcat	gactttatac	gcagtaagta	tggtgaaaac	aaaaaaaaaa	1080
aaaaaaaa						1088

<210> 65

<211> 1256

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (1079)

<223> n equals a,t,g, or c

<400> 65

gggacaagtc	cacgtatatc	gagtcctcga	aggataggcg	ggggaagatt	cctgccaccc	60
tgtgtctctg	gtccagccgg	acgtccatga	ccttgggttg	caggaatctt	cccccgcta	120
tccttcaagg	acaagtccac	gtatatcgag	tcctcgacca	aagtgtatga	tgatatggca	180
ttccggtacc	tgctctggat	cctcttcccc	ctcctgggct	gctatgccgt	ctacagtctt	240
ctgtaccttg	agcacaagg	ctgggtactcc	tgggtgctca	gcatgctcta	cggcttcctg	300
ctgaccttcg	gcttcatcac	catgacgccc	cagctcttca	tcaactacaa	gctcaagtct	360
gtggcccacc	ttccctggcg	catgctcacc	tacaaggccc	tcaacacatt	catcgacgac	420
ctgttcgcct	ttgtcatcaa	gatgcccgtt	atgtaccgga	tcggctgcct	gcgggacgat	480
gtggttttct	tcatctacct	ctaccaacgg	tggatctacc	gcgtcgaccc	cacccgagtc	540
aacgagtttg	gcatgagtgg	agaagacccc	acagctgccc	cccccgtagc	cagaggtccc	600
acagcagcag	gggcccctcac	gcccacacct	gcacccacca	cgaccacccg	caccagggag	660
gaggcctcca	cgcccctgcc	caccaagccc	acccaggggg	ccagctctgc	cagcgagccc	720
caggaagccc	ctccaaagcc	agcagaggac	aagaaaaagg	attagtgcag	actggtcctc	780
acctgtctcg	gctcctggcg	accactaccc	ctgcgtcccc	gccccctcgc	ctccccctcc	840
tgtegccttt	tccctggaca	gatcaggccg	gggcggtggg	aggcccgcct	caggtcaggg	900
cccagcgtgt	gacgtagggg	ccggggcagg	ccagggtttg	tttgtggagg	cgctgtctgt	960
ccctctgtcc	ctctgtgttt	ccagccatct	cgccctgcc	gcccagcacc	actgggaatc	1020
atggtgaagc	tgatgcagcg	ttgccgaggg	ggtgggttgg	gcgggggtgg	ggccgggenc	1080
ccctacggga	tgcccacggc	cgttcatcat	cttgtccctc	gtccccctac	cacactcccc	1140
ctcctagacc	gccgcccttt	aacacagtct	ggatttaata	aattcatatg	ggtgtttaac	1200
ttaaamwmaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaagggc	ggccgc	1256

<210> 66

<211> 1602

<212> DNA

<213> Homo sapiens

<400> 66

ggcacgagaa	aacctgtgga	tttgagtgg	gatgacatta	ttgatgatta	gctgagtgtg	60
gatttgagtt	gggatgacat	tattgatgat	tagctggggg	agtttttgtc	ttcacagtgt	120
tttctcccca	tgtatgaatg	tttctgtct	ctgtctttac	tgaagtcttg	taaagctgtg	180
agtggactta	tgtgcctcct	gctgccgagg	cttgggtctgc	tgtctctcct	accgagtgtg	240
cgatgcttct	gctggattcc	ggtgtactcc	ctcattacct	gccttgctga	gtgctcagtt	300
gttctgcggg	atccagggtt	tgcgggagct	ttccaggtag	acaggcgcca	ggcctgcttc	360
ttccacctgc	gctgggtcctg	cctgctgctc	tgggtgggtgt	cccgggtgag	tgcaggccgc	420
cctctcatag	gcagccctca	tatgatggct	cccagcactt	tctgtcccac	cgtaggggc	480
cctgggacct	gtgcttccag	cgaccagat	gggtgaggct	gtaacagcct	gggccggctg	540
cctctgccct	ttggtgaccg	tgatggggag	gtgtccacaa	agcaccctat	atgttggtct	600
atcctcccgc	caggtctgta	gtctctaccc	ggtcaccctt	tcttctggg	accatgcccc	660
agatcagctg	ctaccctcaa	gctctaccat	tagggggccc	aggcgtgttc	agggaaggct	720
actgggtgcc	accttctccc	ccatcatcac	gctcactgct	ttctcttgat	gggtaaggct	780
tcccgaatga	tgggaactaa	cagtgggtgt	gaagtcgagg	cagacttcat	caggcattta	840
gcattccaagg	ggtggctcat	gagtgcagat	acagtgcagg	ggccggagag	gggctgggct	900
gggctgacct	gggaggctga	ggctgatgga	gagtgggcag	gtagggtgga	gaggcaggaa	960
ggttgggtggc	agccacatgg	ctgagggcta	gagcctggcc	agggagtctg	gagagaggca	1020
gtgggtgggc	tgggggttca	gcctgctgaa	ggggagcact	ggtcagtggc	cataccccat	1080
ggggtgaagg	cctgggcagg	gcccaggggc	agcttcgagg	gtgacctgga	gctgctcagg	1140
aagtgtgatg	gcccagcctg	acctgaccat	tggctggcaa	ggaacgggat	ggagaagttg	1200
tgtcttgggc	cttcagcgag	tgtgacattg	tcattgttgg	gatagcttta	aagatctgat	1260
tgtcttatgac	atgccttgta	gcgctaccag	catcttggca	tttggcagg	ctagtccagc	1320
tcgctgtttg	cacgtcttct	gtcttattcc	tagaagagag	agttcccagc	ttgtcttgatt	1380
tccccccatt	gatgggaggc	tcatacactt	atgggagact	cattttactt	aggccttctg	1440
aggatagttt	cattctgata	gttttttttt	tttttttttt	ttggagactg	agtttccctc	1500
tgctgcccgg	gctggagtgc	agttgtgtga	tctcggtcca	ctgcaagctc	cacctcccag	1560
gttcatgtct	gtgccgaatt	cgatatcaag	cttatcgata	cc		1602

<210> 67

<211> 938

<212> DNA

<213> Homo sapiens

<400> 67

ccacgcgtcc	gctgccagca	gcgcgcagag	ggagggatgg	gggcgggtat	cggcgtaggg	60
ccctcgga	gaacggatat	tgctgtgaca	ccgcggggac	gctctgaagg	ggcgagtgtc	120
ggtgtggcac	cggtgcacgc	tgaaggagcc	ggcggaaccg	ggtggccatg	gggatgtggg	180
catcgctgga	cgctttgtgg	gagatgccgg	ccgagaagcg	tatcttcggg	gccgtgctgc	240
tcttttctg	gacagtgtat	ctttgggaga	ccttcttagc	acagcggcag	agaaggatat	300
ataaaacaac	aactcatgta	ccaccggagt	taggacagat	catggattct	gaaacatttg	360
agaaatctcg	actctatcaa	ctggataaaa	gcactttcag	cttctgggtca	ggactctatt	420
cagagactga	aggcactctt	aatcttctct	ttggagggaat	accttatctc	tggagacttt	480
ctggacgggt	ctgtgggttat	gctggctttg	gaccagaata	tgagatcact	cagtcctctg	540
tgtttctgct	gttggctaca	cttttcagtg	cattgactgg	tgtgccatgg	agtctttata	600
atacttttgt	gataaaaaaa	acatggcttc	aatcaacaga	ctttgggggt	cttcacatgg	660
aaatataata	aaaaataaaa	tctagtttaa	tactgcatta	tttattttcc	taaggctaaa	720
gaggagcagt	cctatgcttt	tattcagcat	cctttatctg	tgacttcatg	ctctgataac	780
tgcttttct	tccttctgtg	cctttgaata	caaatttcag	ttctgcaaaa	gtgaaacatt	840
aaacattgcc	aacgcaaat	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	900
aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa			938

<210> 68

<211> 1585

<212> DNA

<213> Homo sapiens

<220>
 <221> SITE
 <222> (904)
 <223> n equals a,t,g, or c

<400> 68
 ggggttgctgt tgggggtgtgt cgrgaggacg tcatgggaat tactgatcgt tcaaaaatgt 60
 cccagatgt gggcatctgg gcgatttatt ggagtgtgc tggctattgg cccttgatag 120
 gcttccctgg aactccsacc cagcaagagc cagctctcca ccgagtggg gtttacctgg 180
 atcgtgggac tgggaatgtc tccttctaca gcgctgtgga cggagtgcac ctgcacacct 240
 ttctctgttc ttctgtctca cgcctccggc cattttttt gggtgagtc attagcatct 300
 ttagtcattc caccagtgc tgataggaaa tgaggctttt cttccctga ccaaaactcc 360
 ttccctgtag tccagtgcg ggacacacat ccctgggccc tcttctgccc ttcatgtctc 420
 tatcctggat ggtccatctt ctgggtctcc ctaacggtac cgtttggtat ctgccctttg 480
 tgtgtctcac aagaggcagt cccatgggag gtgggtctgg ccaatggaga tgggacagga 540
 aattttccaa gacgcttttg ggaaatcttt ttgtagcttt taaagagatg tgcgggggaaag 600
 acatatggat gttagcagcc atattggaac tgagaacaca ggaatggagg aaaggtagaa 660
 gaaacagagt tttgttgcc gttgtgaaat tggtgaaatt tccttcatgc caagcttctt 720
 gttatatgag ataattacgc cttattgta taagacaatt ttagttgtat ttggttactt 780
 gcagcctgaa gtaccgtaac trcactaaag ggacgtagtg tgaacatccc gcagtatagg 840
 cttaagtcac tttgtgaaa ttgacaaaag gcatagatct ttttctatcc agtcaggcat 900
 tgcntattct ttccagtaac tactgattcc cccacttttc tgtcttagaa aattgtggga 960
 atccccctc actctgccta tgttgactc tctctctcc caaccataac tctgccctca 1020
 gctattaact gtgctgtgta tttatcaagt tggtagttg tatgtacagt gttattcatg 1080
 tcatcatgag aatgttgaac gcttgttaa tatcttttg tagcctcttc atatgctgtt 1140
 gcatatgact ctcatcaca ctcagtgaga tggaaagtca aatcctattt gtacaaatga 1200
 gaaaactgaa ctcttttagag taactagctc agtattggcc agctggtaaa tggcagtggt 1260
 gggattaaaa tccagttctt atctactctc cctttattca gaagcattta ttggatgttg 1320
 atctttgttt caggttttg tttgttact tttttatact gtgtatattt tctcagtcct 1380
 acccttctgc tctagattgt ctggactcag gagattgtgg cagttactgg atagttattt 1440
 ttaagataat gattgctttt ctctgtttat ataagtcag tgtacttatt gtagaaagtt 1500
 tgtaagatgc aaaaagtata aaaattaaag ttatgcacta ctaaaaaaa aaaaaaaa 1560
 aaaaaaaa aaaaaggcg gccgc 1585

<210> 69
 <211> 1676
 <212> DNA
 <213> Homo sapiens

<400> 69
 ggcacgaggg gacttcagaa ccacagaact gagatgataa atgagtgggtg tttcaagttg 60
 ctaagtttgt ggtcatttgc ttacagtaat tgtaaactaa tacacaagtg taagtttgtt 120
 ttcttaaga agaaaaaac ggggaaggag gtaagtgtta aaggatcaaa actctgacaa 180
 aaggctggtt gcagaacatg acaggttgtt gcaactggaaa ctatttgtca tgcaagttta 240
 tgttaaaata agtagctttt gaggactttc atttttggtc ttgtaaacat gccattta 300
 attgtccmac tgataatact ttttgcaaac agaaactgtt aaaaccttta aagcaaatat 360
 tactgtagag aagaagtaat gtgttatgaa actgtgagga tactaagaag gatcctactt 420
 aagtttcttc agcataaata aacttgagcg ttctgaccac tgttactgag aatgaaatta 480
 tttcttaatc acttttaatg aggtaaaatt tacatacgat aaaatgcacc aattttaaag 540
 tatagtttaa tgagcttgca cagatgtaaa tatctgttta acttctactt aatcaagata 600
 tagaatattt ccacaatgcc aaaattgcca ttgacccctt tcccccttct tcaccaact 660
 gcagacccca ggtcaccacc aacctactct tgcctaatat agatttaagt tgatgtgtct 720
 tttctagagt tttatgtcaa tagaattgta cactatgcac tcttccatgc ctggctttct 780
 ttgctcagca gragggtgtt agattaattc agtagttcat ttctttctag taatgaatag 840
 gatcacatta tacattatac cacagagtgt gcatccatta ctttgtkgat tgatatttgg 900

gtcatttcca	ggttttggct	attgtgaata	aaactgcctt	gactattcct	gwacaagtct	960
ttgtattaag	gaacatacgt	tttattttct	cttgaggaag	ttcctagcaa	taagattgct	1020
gggtcatatg	gtaggatat	atttagcttt	aaaagcaact	aagtgccttc	caaagtgact	1080
gtacaattta	acattcctac	ctgaaatgta	agagaattcc	agttgctcca	cattcttgct	1140
aacccttggg	agcatcagtc	tctttaagaa	ttctaattga	tatgtaatat	ggactatagg	1200
tttaatttgc	atctctctgt	tgactaatga	tggtgcacaa	cttttcatat	gtctatcaac	1260
cattcttgca	tcttctttta	tgaatgtct	gttcaaatca	tttgtccact	ttttattgtg	1320
tcattttatt	cagttgtaag	agttctttac	atattctgga	aacaagtcct	ctgtcacata	1380
tataggtaact	ttgaaaatct	gtgctttgcc	tttacatttt	tttaattgga	actttttaag	1440
agtagatagt	tttggttttg	atgaaattca	acttatcagt	ttttcagtta	tagtatgtat	1500
ttttatgacc	catctaagaa	gcattctgtct	acccagagtt	gcaaagatat	cccttttctt	1560
actagaaata	tttagtttt	atttaccatt	gcttctatga	tacattttta	gttaattttt	1620
gtgtattaaa	tgaataaaaa	gttgaagttc	aaaaaaaaaa	aaaaaaaaact	cgtagg	1676

<210> 70

<211> 1344

<212> DNA

<213> Homo sapiens

<400> 70

ggcacgagcg	ccagataact	caactttccc	attggctacc	tttgggtcag	gtgatctcca	60
ctagacctat	cgcctatgcc	tgatgggtgg	tcacatgggt	caaagtgtgc	ctgagagctt	120
agtggattag	ggatgtggct	gggtcatggt	ttgacgtccc	tgctgctgag	cccttacggg	180
tcaggctggg	agaaggtacc	atgttgtgtg	actggctcatt	tgaggtcttg	cagctgttgc	240
ttgctgggct	tggcaggtgt	tcaaagtgac	catttttctg	aagggttttt	ttctgagtat	300
tcctcagatg	tactccctcg	gggcccagcg	tctttccctc	cacagggcga	tgcttcccta	360
cttgcttgtg	aatgtttcct	tcattctccag	gttgtctggg	gacaattctg	tcttttggag	420
ggctgggcag	gatttacaga	gggtccatg	ccagctcctt	cctgccgggt	ccactctggt	480
tgtagggtaa	acacctgccc	attcatgtcc	tagtgttgat	agaataatca	ttttctttca	540
gtacagtttc	cttttttttt	ttttttgccc	cagcttttta	gatgtagcac	ttaatgccag	600
ttctcgagct	ccccaaactt	aagggtacaca	ggtcaacaag	cagtaggtct	ttggaagctc	660
gctctctcac	atggtataag	gtgaggggga	cacatggaat	gtaaacctcc	aaactaatta	720
tgggggaaaa	aggaatgaga	aaaacaaaca	caagaaggca	aaacaaaac	acctggttca	780
attaaaaaca	acaacaaagc	aaaacaaaaa	aaacccaaaa	ccaaaccaca	caataaatga	840
gaaaaaaatt	acataaaaca	gatgacatgt	atagaaataa	aaaaatagaa	tggaggaact	900
aattccaaat	atataagtaa	tcctaagtgg	actaaatttg	caaagtataa	ggcaaacatt	960
acactgggatt	aaaaaacaaa	tcagggaat	acagttcaca	gctgatacat	tcaaaatata	1020
aaaacagata	aaatataaat	aaaatcggta	cagtgggtgc	tgctgtaat	cccagctagt	1080
tgggaggctg	aggcaggaag	accccttgag	cttaggagtt	agagaacagc	ttgggcaaca	1140
aagcgagact	gtctctaaaa	aaatacacac	acacacacac	acacacacac	acacacacac	1200
acacacacac	acacacacac	acacacacac	acacacacac	agactgaaat	ggaaatgagc	1260
cgagatcgcg	ccactgcact	ccagcctggc	aacagagtga	gactctgtct	caaaaaaaaa	1320
aaaaaaaaaa	aaaaaaaaaa	aaaa				1344

<210> 71

<211> 1474

<212> DNA

<213> Homo sapiens

<400> 71

ggcacgagct	aaaatatgct	taaagtaaga	tgtttctatg	ttatgtggtt	atgttatcaa	60
taatatattg	ttgatcttca	catattttat	atgcatatat	atatcaagaa	gttatatata	120
tataactcaa	gaaactcaag	ttatatatat	atgtcaagaa	atgtatgatt	attttggaga	180
gaatgggccc	aatgtgaaa	aagatataaa	aaaaacaaaa	aaaacaaaaa	aaaaacacta	240
ttttcccta	cgaaatatac	tgtatatattc	aaaagaagaa	aagttaaaag	atatttgaag	300

attgcagggg	caaaacaaaa	acctaccagg	gctcagcact	taagcacttt	tcccacatcc	360
agggtcaaag	cagcagtaac	tacatggact	tttaagtgcg	gtatgaaatg	caacattacg	420
cttacaaaca	gtactgtcaa	acctcaaagt	ctttctttct	tcagatgctt	tttcgtgtac	480
atgatactag	tagacacttt	tctctttata	tttactgata	gtgaaaatca	tacgcaataa	540
aattattgatg	tttgaaggca	gtggtcacca	attgggttaa	aaactatgaa	atgtaaactg	600
aattgttata	tctctatcct	ttttgttttt	ctctgtgttt	ttaatgtatg	gaataaatct	660
cataaataga	aagaaaaata	atctagaaat	ttttcaaagc	tagtactctt	tctccttata	720
aatgtacaca	attttaaatct	ttttacaaat	ttattttaact	gtacctactg	tacttattgt	780
agattcaatg	acgcagttaa	gtcatcaccc	aaggattttat	gaatttgaga	ttactgacct	840
gttttcttca	tattgcattc	acatcaatat	ttgtgaattt	gttggttcagc	ttttcattca	900
aacaaaaaat	attccctcaa	gaaagctccc	tttttatcat	aaacatttca	acttacccaa	960
cattagaaca	agtctgccat	gttaaaaaata	atttaaagac	ttatctctga	aaacgggtatc	1020
cagaaacgca	ggtgttccca	gtaatgtagc	ttcaaaaaata	aaatgtgcta	tttatatgac	1080
tgaattcat	aacttttggg	aggggtatatt	tatgacagca	taaaaaataa	attctgtgct	1140
ataaagaaga	tccaacaaat	taaccatata	agcacagaaa	gtagagaaac	acagttattg	1200
aatctactct	tgtcattaac	attttcaaaa	aacaaaatgc	atattgtaat	atttgggtaca	1260
tgacacttgc	atgttgatat	gcctatatata	ttacaaaagta	ttcaatgtgt	acttagcggc	1320
gcttaaaata	tgtcatgtac	aactcttata	aacattttta	cagggttccc	atttgcactt	1380
catctttcag	taaagtcttg	tcagaaaaaa	attgtctgat	aaatatggaa	aaataaaatt	1440
tgaatttttag	ttaaaaaaa	aaaaaaaaaa	aaaa			1474

<210> 72

<211> 2012

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (1468)

<223> n equals a,t,g, or c

<400> 72

aatttttttaa	aagcaatttt	aggttcatag	caaaatttag	cagaaagaac	agagttctta	60
tatacccttt	ctgcttttcc	ccaagcctct	tgcactaata	acgttctgta	ccagagtgggt	120
atgtttgtta	caatcagtga	acctagactg	gacacctcat	tatcacccaa	gaccaagggt	180
tacgttaagg	ttcgctctta	qtqttgtata	ttctatgagt	ttgacaaatg	tttaattgaca	240
tgtatccacc	atgatagtat	catacagaat	agtttccactg	ccttccctctg	tgctctgcct	300
gttcatccct	cccttccctg	tgatcttcta	ttgtctccat	agttttgtct	tttcccaaat	360
gctgtatagt	tggaattatc	atgtaacctt	tcaaatggcc	ttttcactta	taatattgct	420
ttaagggttt	ctacagtctt	kgcaggcttg	atagctcatt	tttcttttagt	gctgaatact	480
attccattgt	ttggatgcat	tacagtttat	tcattcacct	actgaaggaa	atcttgccctg	540
cttccaagtt	ttgtcaataa	agctgctttg	taaacatcca	tgtgcagggt	tttgtgtgga	600
cataagtttt	caactcattt	gggtaaatac	caagcagcac	aattgcccgtg	ttgtatatga	660
agagaatggt	tagtttcata	agacatcacc	aagctgtctt	acaaagtggc	tgtactattt	720
tgcacttcca	ccagcaatga	atgagagttc	ttgttggttc	cacatactcg	tcagcatttg	780
atgatatcag	tgttttagat	tttgaccatt	taatagggtg	gtagtggcat	ctccttggtg	840
ctttaacttg	taattctcta	atgacttata	atgttgagca	tcttttcata	tacttatttg	900
ccatttgtgt	tctctcttcc	ttttcttttt	tctttttctt	tttttttttt	tttttttgag	960
acaggggtatc	actctgtctc	tcggactgga	gtgcagtggg	acggtcacag	ctcgtctgcag	1020
cctcgatgtc	ccaggtcag	gtgatccctc	catctcggcc	tcctgagtag	ctggaactac	1080
aggcacatgc	caccatgcct	agctaatttt	ttatattttt	agaagagaca	ggtttcgtca	1140
tgtagtcag	gctgggtctca	aactcgtggg	ctcaagtgat	ccaccacact	tggcctccga	1200
aagtgccggg	attgcaggca	tgagccackg	tgcccagcca	atcaaatgag	aaaagggcag	1260
tctcttcaac	aaatgggtgt	gggaaaattg	aatatccaca	ttcaaaaagaa	tgaatttaga	1320
tgcttacctt	aaactgtata	caaaaattac	ctcaaaaag	atcgaagtcc	taaatgtaag	1380
acctaaaact	ataaaaactct	tagataaaaa	cataggagga	cattggattt	ggcaatggat	1440

ttcttggg	cg	tgacaccaa	aacaggcnac	mgaagtaaaa	atagatggat	gacatcaaaa	1500
tttaaaactt	ctgtgtatca	aagaacacat	tcaacagagt	gaaaaggccc	ataaaatggg		1560
tgaaaatatt	tgcaaatgat	atatcctgta	agaggcta	atccagaata	cataaggaaa		1620
tcctaccatt	caataacaaa	aaacaacctg	attttaaagt	gaagaagtca	agaaaaataag		1680
atagactaca	tactgggaga	aaatatttgc	aaaagatata	tctaataaag	gactgttatc		1740
caaaatatac	aaaagaactc	taaacgataa	gaagacaaat	agcctgaata	aaatacagca		1800
aaagacttga	acacataccc	cactaaagaa	gattcacaar	rttgggtgca	gtggttcgcg		1860
cctgtaatcc	taacactttt	ggaggccaag	gtggggagaat	cgcttgagcc	tgggaggtcg		1920
gggctgcagc	gggctgtgat	tgtgmcgctg	cactccggcc	tgggcaacag	agggagaccg		1980
tgtctgaaaa	aaaaaaaaaa	aaaaacctcg	ag				2012

<210> 73

<211> 1267

<212> DNA

<213> Homo sapiens

<400> 73

ggcacgagct	cactctagct	gctatgataa	acaagaagac	cagtgaggaa	gttttgaaac	60
tctcctctgc	aagagaatgg	tggccagcca	ggcgtgggtg	ttgtcgaatc	tgtggcacct	120
gtgggaagtg	ggctcagccc	agggactgcc	tttggatccc	ccagcattgg	caccatacct	180
accctgggcc	ctaagggtggc	catgcttctc	tggatttgct	tccctagcag	gggctctagt	240
gcttgeccac	tccctgcccc	cagcatggcc	tgggagcagc	tgagactggg	ggccagctca	300
tcccgcgtcg	attcctggaa	gtgttatcag	tgctgttat	ggaggctgga	cccatgagt	360
gcagccttcc	ctggcagctg	ggctgacctg	tctgcttttc	cattgctcgc	tggttttgtt	420
cactgtaggg	cgtaggggtg	gagtagctgc	tggcctccaa	gtccatagct	actcatgttg	480
tacgctgttc	acagggacct	ctaaggatgt	acatcatcac	acattcacac	acatgcacca	540
ggatttgctt	tttttggcag	cttttccctt	cctggcttcc	tttttgagt	gtggaagtaa	600
aataaaaaagc	aactagggct	gggcacagt	gtcacgcct	gtaatcccag	tacttttaga	660
ggctgaggcg	ggcagattac	ttgaggtcag	gagtttgaga	ccagcctggc	caacatgggt	720
aaaccccgtc	tctactaaaa	atacaaaaaat	tagctgggag	tgggtggtgca	cccctgtagt	780
cccagctact	cgggaggcta	aggcaggaga	atcacttcaa	cctgggaggc	ggagggttgca	840
gtgagctgag	atcacaccac	agcactccag	cctgggtgac	agagccagac	tgtgtctcaa	900
aacaacaaaa	caaacaacaa	caacaacaaa	aaactgggct	ctgggtgggt	ggaggaggag	960
ggaagggaagc	aagtaccagg	gtaagcagga	tggatggatg	gctctcccc	agagggggcgg	1020
cagcacacag	agttctggag	tcagactcag	tagagggcca	gttttgactc	cactgccaac	1080
cacttggtg	actccaggca	ggttacatca	ctgatgaaaq	cctcagtttc	cttgtctata	1140
aattgggggt	acaagcgatg	aaaagagggt	caacatcact	aatcactagg	gaaatgcaaa	1200
tgaaaatcat	aaggagggtac	cactttgtac	ccttaaaaaat	agttactaca	aaaaaaaaaa	1260
aaaaaaa						1267

<210> 74

<211> 1748

<212> DNA

<213> Homo sapiens

<400> 74

ggcacgagta	aagacaaaat	aaattcttct	gtccacttat	ttacctaa	tacacttget	60
tccttggaag	tcataggcat	ccacatatct	tcagccacaa	cttggatatt	ctaataataat	120
tcttattttt	gaactcctaa	actttctggt	agaaatcttc	agttgaaaat	atcctggcaa	180
gtaaaattag	aaactcccag	aatgtactt	atttctatta	tgttgtttta	tttctgaaca	240
ttgtgcccac	cattcttttc	cacatacttg	cccaaattgg	aaaactaggg	ttctaagttt	300
ccccctccat	ccatgcccac	atttaattca	ccctaataat	acctgacatc	tttcaagttc	360
attttctact	atctatcccc	acgcaggcca	tatctgggtt	gaagcttcat	tatctctata	420
gattaaaaac	aaaaacaaaa	tgcatacaag	caaaacaaat	aacatacaaa	caaaacccac	480
ctaactcatc	tttatgtagt	cagtcctccc	tcaatagttt	ggccaaactt	cctaaaccga	540

aatctgattg	tgactatccc	cttctaaatg	tatttaaatca	gcatacccct	tcaataaatc	600
cattaaccgt	tcttggtatc	caagacagtt	tgatcatctgt	cttggataac	aagttgcaga	660
ctccatccaa	tgccattttc	ccctagaaat	atagataatg	gcactatagg	aacaatgatc	720
tccacatgcc	tcatgcattg	gtaatttttt	taacctttgc	tagaaatgtt	ctgctccact	780
ctactcatcc	accacccatt	ctactccacc	ttacactacc	tcttttccca	tttagatctt	840
ccattctata	tctccttgca	tgaaattgtc	catacctgct	tagtactcat	ctcattatct	900
tgatcattgt	cgcataatctg	ttcatatgat	ttcttatgga	agttattaag	tattttgatt	960
tctgtttcag	tcagattttc	aacagagaag	tagaaccagt	agaaaatata	tcttaagata	1020
cttattggag	ggaattaact	tacatgggtg	tggggaaccg	catagccgac	ctaaaattta	1080
tatgggtgac	tatcaaaaaa	gacaggctgg	aactcttagg	cacaggcaga	agttgcagtt	1140
cacagggtgaa	atttggtctt	tatcctggaa	gacctgggctc	tgctcttttag	atttagcagc	1200
tgactgaatc	aagtccacct	agattaccta	ggataatctt	gtttacgatt	atgattatca	1260
ctaccagtta	tcaactgatt	ttgaacttca	ttcacatcta	caaaatacct	tcataaggaa	1320
atctagatca	gtgttggtg	aaataactat	cagctgtagc	ctagccatgt	tgacctcatca	1380
aaagaccatc	acaattgctg	atataacttt	aataaaatct	gcaacatttt	cagatgggaag	1440
aatttgagaaa	aggggaagcg	gctgactttt	catttttagaa	tttattatgc	attaacttaa	1500
agtaagtaat	aattatgtag	gtgatcattt	tgatatttta	acctacttaa	tttagaaaaat	1560
cattttaaata	catttttgtt	aagactacaa	aatgatcttg	ggtaaaaaaa	aattttacca	1620
aatatcaaga	tcacaataat	cactttaaata	agttacatat	gtaactaacc	tgacacaatgt	1680
gtacatgtac	cctaaaactt	aaagtataat	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	1740
aaaaaaaaaa						1748

<210> 75

<211> 1570

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (7)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (8)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (10)

<223> n equals a,t,g, or c

<400> 75

taccggnnan	ggaattcccc	ggtcgaccca	cgcgtccgat	tttatgctat	ccagtwttta	60
cgtacccatg	tgtaacatt	tcatatatcc	agttcttttg	gtgctgggtc	actttttttc	120
atttattcaa	attcagaaaa	atacgacgg	atctaagtgt	aaattaacca	gaaaccctgg	180
aacattcata	tcttgagagg	gaatttgcta	ctcttttact	tttgggattt	cattataaaa	240
taggctcatt	ttatacatat	gtcctgtggt	ctgtcttcca	gagtgtgctg	taatcataag	300
tctctagcaa	agagtggagg	gtggagggtg	tgtagaactc	cactcagcct	catggataca	360
tctacaagtc	tctcaacca	tcctgttaca	gttctccatg	agagtcacct	cacacttgga	420
aaagaaggag	gtttaaagca	gagtattgtt	ggcagccagc	tatgcccata	ctcttgtaaa	480
taactttcca	acacacccat	ttccttgcc	taagggtgcag	gttcctactt	ctacaaaatt	540
tcaaatagatt	gcagggtcaaa	aacctttgga	gttactcaca	aataataata	aaattttcaa	600
taactagggt	ctacttgtcc	atcaagatga	cattaccagt	ggaaccaccc	tggttaatttt	660
aaagtagctt	cagaacccaa	caaaaattat	gtaagcctgg	ttaaatagtcc	ttttttcttc	720
ttgcctctaa	taraatcagg	atcttcggcc	ttgatactaa	atatgtgtat	atttagttat	780

gatggtactt	gtagatgctc	acatttccag	ttccaaactc	gcccgtactt	tttatgtgct	840
tcaaaatatt	ggacacattt	ctgttaatat	atgatttctg	tatccacaaa	ccgctgtttg	900
cttatgctga	gtcaattttag	aagttaattt	ccaatctagt	ctcaactgca	atgcatttaa	960
aataggtaaa	acaagtaaat	gagtttgga	cattttgaga	ttaatgttac	tgcccacttg	1020
actgtcaatt	tcaaatggct	cctaattgca	ccaaaattat	aaccaatgta	gcatgtgtag	1080
gaaaggtatt	tttaattatt	taaaatcatt	gtgtatatta	cagcagtatg	aggaatgcct	1140
ggctaaagag	gatttttttaa	aagatgaaga	atggttttgc	ttgtattata	taggcttact	1200
gagtttgtga	gcagcataaa	aacaatcatt	ccttaattct	tcattgtgga	actgaaatat	1260
ttctgkaatg	catttttttaa	aaggagactc	ctagacacag	cygtaatagg	ggaatagaaa	1320
tgtagagctt	ttctcataac	acaaaaccag	aaaataattt	tcagcacttt	gacttactct	1380
atgtaataag	gaaaaaaaatt	gtttccacaa	agttgaacta	tgtagtaata	tttggttaaca	1440
tatggcatgg	ccactttata	tcacagaatg	tgtgtcaagt	tgcaaagcat	acttgggcca	1500
tagtagacac	taagaattaa	aactcttaaa	tcagtgaaac	aaaaaaaaaa	aaaaaaaaaa	1560
gggcggccgc						1570

<210> 76

<211> 524

<212> DNA

<213> Homo sapiens

<400> 76

ggcacgagcg	gcacgagcgg	cacgagacgg	cctcccctcc	gccctgcggt	cccgcgcgga	60
tgcagatctg	tggatccagc	gtagcatctg	tagcagctgg	gacatcattc	caggttttgg	120
gcccgggtg	ttggcaacaa	ctggatctga	agatggcagt	caggggtgctt	tgggggtggtc	180
tcagcctgct	ccgagtgcctg	tgggtgtctcc	ttccgcagac	gggctatgtg	caccagatg	240
agttcttcca	gtcccctgag	gtgatggcag	gtaaaactcc	gcatgtgtgg	ctgagacaag	300
ctgcagcaga	gtctgcttga	gaagctgacg	ggagactttg	tggggaggga	gtagcatgtc	360
tgggtagatg	agtagtaaat	ccacaagcag	agcagcagcc	tctctctctg	gggtaagaac	420
ttggaagtgg	ggacttcata	tctccttccc	gagtgggtgac	actgaccttc	tgggtaatgc	480
taataaacca	ccagtctctt	tgatgtaaaa	aaaaaaaaaa	aaaa		524

<210> 77

<211> 1306

<212> DNA

<213> Homo sapiens

<400> 77

gaagctcgaa	attaaccctc	actaaaggga	acaaaagctg	gagctccacc	gcggtgggcg	60
ggcgcgtcta	gaactagtgg	atcccccg	ctgcaggaat	tcggcasrrg	gcaagctgag	120
atcttcaacg	cttctacaa	gaagtaccta	gataggagat	gggagggaaga	gccactcagg	180
accaagactc	tgccagcagc	tctccttgcc	agggagtcta	catggctctg	aactcgtgtg	240
tctttctttc	ccagtacggc	caccttctat	ttccttcttc	ctagctgcct	atttgcaatg	300
ccaccggaag	tcaagggcc	ctcaggcaag	gagaatagca	cttcataaag	agaaggatga	360
tgaccccgag	ggtgtgtggc	cctgtgctgc	gcccattgca	gtctctcagc	tcagctgctc	420
ctcctcctac	ctggtgctgg	cctgcgagga	atggtgtgct	cacgctgtgg	gacctggcca	480
aaggattccc	tcttggggtc	gctgctcttc	ctcaaggatg	tttctgccaa	agcattcact	540
tcctaaaata	tttctcggtc	cacaaaggac	agaatatgta	tectgaaggt	caagtgaat	600
cccaaataaa	atgtgtgtgtg	ctgtgcacag	acgcctccct	ccatctgggtg	gaggctagcg	660
ggacccaagg	accaccatc	agtgtgcttg	ttgagatgtg	ctcatctttt	ccaagaatgg	720
ctctgtgtgc	cttatggatg	tggccaagcg	tgaatcatc	tgtgcctttg	ccctccggg	780
agcctttcct	ctggaggtcc	cctggaagcc	agtgcttctg	gtgtctccag	accatccatg	840
tttctgtctc	cgaggcctgc	ccactcctgg	aaaatatctc	aaaaaattgt	accattcctc	900
aaagggactt	ggataacatg	gccttcccc	aagcactgcc	actggagaag	agatgtgagc	960
gtttcctcca	gaagagctat	cggaaagctg	agaagaacct	agagaaggag	gaggagcact	1020
gggcccggct	tcagaggtac	tccttgtcgc	tccagagaga	gaacttcaag	aagtgaggct	1080

gccaccgccc	tgggatctct	gaaaaggagg	tttcagccac	gaggcagctg	cttccaggac	1140
actgaggcca	agagaaatgt	aacagarcca	cagctccaca	ggcctgcact	cggagtctgg	1200
ggcctctgca	gaaccagcaa	ggggaaaagt	ataatctggg	ggaccttcaa	ccactaagcc	1260
tcttgtcaga	accctcaggc	agggcagatg	tgtcaccaaa	taaaac		1306

<210> 78

<211> 1479

<212> DNA

<213> Homo sapiens

<400> 78

acgsgaagct	cgaaattaac	cctcactgaa	agggaaacaaa	agctggagct	ccaccgcggt	60
ggcgggccc	tctagaacta	gtggatcccc	cgggctgcag	gaattcggca	cgagcttggt	120
tcatgccctt	ccaattacct	ctccagcttt	tactacttag	gttaatttgt	gaattctttc	180
ttgcccctgc	cttaaattgt	aatctcactg	ggactgtgat	tttctttaca	ctcatgattt	240
ctttacagct	catgattttc	tttacctgc	agtttgctga	tggcttccaa	attggtgttg	300
atcttcagct	ttcagagttg	aatattctgt	aggtggcttc	atatgtcatt	cattccgtct	360
tctcagttta	ttcttaaata	tgacccaact	ccagccttta	tcccaattta	tactacctct	420
gtcctgattc	aaggcttcat	tatctgtcac	ctaaatgtac	ttgaagctgt	agtgccctgc	480
ttcattcttg	atctcttcaa	atctgctttt	ttataaggcc	agcaggatgc	tttggcaaaa	540
atactgatct	tagtgactgg	tctttttaaa	agcagaaatt	atcagatctc	tgataagaat	600
ccccaagagt	aggatacctt	gaagtctatt	ttgtatacaa	gttcattcca	tacctgcaga	660
tctggtttta	atcagattga	gaaacttgca	gtctaccaga	attaagtttg	agccccctctg	720
cctggtgcat	aaagtctctc	atgacctggg	tttgctttac	ctgtagtcca	catttcccac	780
tatggcctaa	tttgagcag	tattgaagaa	cacactctct	tactgtcctg	tccccctacc	840
tgctgtatcc	tttgcataga	ataagtgcc	ttctcttcca	gtctacataa	cgtaccttca	900
cttttttttg	tggcagctgt	tattgcactt	tgcaacattt	gccaaaacgt	ttctgactcc	960
ccagcctggt	ttatagtaac	actcatatgg	acaccttgaa	aattttataa	atattatttc	1020
attggatgag	aaaatgaatc	aatttttaata	aaaataattt	acaccaatgt	taacattaaa	1080
gattgcacta	gagttcaaaa	ttagtatatt	tttatgttta	aagaaatttt	aggccaagtg	1140
cagtggcctg	agcctgtaat	cccagcactt	tgggagactg	aggcaagtgg	atcacttgag	1200
gtcgggaagt	tcaagaccag	cctggcta	atggcaaaac	cccactctct	ctaaaaatgc	1260
aaaaattaac	aatgcttgat	ggcaatgcct	gtaactgcag	ctatttggga	ggcagaggga	1320
ggagagtcac	ttgaacttgg	gagttgccgt	gagctgagat	agcgccagtg	cactccagcc	1380
tgagtaactg	agactcctat	ctcaaaaaaa	aaaaaaaaaa	aaaaactcga	ggggggggccc	1440
gggtacccaa	ttgcgcctca	tcagtgcagt	cgtattaca			1479

<210> 79

<211> 1794

<212> DNA

<213> Homo sapiens

<400> 79

ggcacgagaa	gcatttagta	ggatttttaa	gaaacttgag	aactgttaca	taaggtgatg	60
aattgggcat	agcatgtaa	atttatattt	agcaaggaaa	tgatctctgg	tgttttaata	120
ttcaacttga	ttgcttcctc	ttgggttctg	tgtttccac	tgtgtgacct	gagctgtcag	180
aaaaccttaa	gaattttctt	tgcatcattt	ttccatgcag	tttgtgtaca	tgtctcatgt	240
acctcgtggc	agccactggg	tttgttcata	aaatggtggg	ttgtgggatg	ctctcctgca	300
gtctccctct	aattaaagag	gttaaattgc	cgtttgtctc	gccttttagtt	cctttccaca	360
gcttcctagg	ctcttaaaaa	ttagcactat	attcctttca	gattaaaaaa	aaaacaaaaa	420
caaaaacctg	tttgcgtgct	ttactgctgt	ggcttgtctc	agagcaaatc	tgaacaaact	480
gattgaaagg	ggtgtttggt	ggctggtggt	ctctttgact	aaagaggctt	acatgtactg	540
tggtacagtc	tgcttactta	aaaggtgagg	cttgaattaa	aatacagcca	gatagaagcc	600
agactcta	caaatgaggt	gattagatca	atgaatgaag	agaggagagg	agtcagggtg	660
tgcttttccc	tggctgttga	atagctgatg	ttccagattg	ccctacagtg	ttgtgttagg	720

gcacccagga	gggatacttt	tcaggcttag	gtacacctca	gtcttttaaaa	tgaggaatta	780
ggacacattc	atgtgtgtgt	ccctaactctg	ctcctgagaa	gagaagtgc	atcagggtct	840
tattttgtga	ccactgactt	gcacactgag	acaaaagggc	catctgcaag	ctgaaaatag	900
tggattcctt	aaataaaaac	tattcacatt	tgatgggtgtg	gtagttttta	taaaatgttc	960
aagtgtcaag	ttcattttca	tttataatct	gagacagttt	tataagtcac	ctccctgggg	1020
gtaaaaatgc	atgtttctgtc	ctcatagtga	gacacatctt	ctgcttagag	tctagaaagc	1080
tctaagaaag	atttatgcca	tctgtgcagc	tgccattttt	atagtaaaat	tttttttact	1140
ttgctccaag	tttaagttat	ctcatgacaa	actttcttga	aagaggcatt	cactattatt	1200
ataggaagta	tactttcttta	ttgaaaagga	gataatgtat	caggtaactt	attaaagtat	1260
tttctcaaag	tttagtatct	tttagaatac	agtgcctcaa	tacaataata	aattatttgt	1320
aaataataga	atgaattcat	tttagaattt	aaatgatgct	aataaaatag	accattatct	1380
taaaagttta	actaatcttag	aatcaaccct	gggtgaaaat	aaagccttaa	gctgtttttt	1440
tgggaagactt	taaatccttt	atggctaaga	gatgacagac	agggccgagt	gcggtggctc	1500
atgcctgtaa	tcccagcact	ttgggaggcc	gaggcgggcg	gatcacgagg	tcaggaaatc	1560
aagaccatcc	tggctaacac	ggtgaaaccc	tgtctctact	aaaaaataca	aaaaaaatta	1620
gccgggctgt	gtggcgggcg	cctgtagtcc	cagctactca	ggaggctgaa	gcaggagcat	1680
ggtgtgaacc	caggaggcag	agcttgacgt	gagctgagat	cacaccactg	cactccagcc	1740
tgggcaacag	agcgagagag	tgagactctg	tctcaaaaaa	aaaaaaaaaa	aaaa	1794

<210> 80

<211> 1280

<212> DNA

<213> Homo sapiens

<400> 80

ggcagagta	taaaggcccc	tccaccagc	ctctagccca	tttctttctc	tggtttttgc	60
aggattcctc	attctccctg	aggtcttaac	tatcacatca	tgcaggttct	gccactgctg	120
ttatcactgc	tgctgctgct	gctgctgctg	tcagctagct	ttgtgacttt	cagcaccccc	180
acttccagca	gaaattctag	ctgccctgat	tgtgagagtc	tgaacaccgg	tcttccatcc	240
ctgatgatgt	ttggtggatc	tctgctcaaa	tgggttcaga	acacacacgg	ggtggaatca	300
ctctgtctct	ctgccaaggt	gcgcctgctt	ccaccagccc	taggggttct	gttcccaaga	360
ctacaccctg	gcactctgac	ccttgtcttc	cttttaattc	ccttcctcac	agtgtcttct	420
tccacatctg	acgtttcttag	ctcttttagag	tccccaaaac	tatctgttac	catatttcat	480
tattgttaac	tctaaagatt	ttggcatcaa	acaccctgca	tttgaatgct	agctgtgtca	540
cacatcagat	gctttacttt	ggcaaatcat	agaactttct	gtcaataggg	ataataatgg	600
tacctatatt	gtaatatgtt	gagtgttact	tggataataa	agtacatagc	acagtatttg	660
gcacatagta	attgctcaac	aataccaatt	gttattatta	ttagactgtg	ccctctaaat	720
tattttgtcta	cggattatga	tctgtataaa	tgacttatca	attaagaaga	ccacagggtat	780
gcagagtctc	atcactcata	caagactgat	gtcaattaac	taagagaagt	ttcgtcacta	840
accaggagtt	tcacatcata	gttccacact	ttgcttctac	tcccaatatg	gcttggtgac	900
ttttcactct	ctttaccctg	ttttctttct	atgggtccca	gggctatcac	ttttctttat	960
tttgggttaat	acatatagct	gtacactgac	ccagctctcca	tgaaaaatac	tgtcatatac	1020
tccctctttc	cctctttccc	taatatcacc	atctcataga	gatcaaaact	acattttctg	1080
gcaccattat	tctttttata	aaatacttta	cttttaaatt	tttaccacac	tacgtctatg	1140
ttatttttagc	tagctaagct	gctataacaa	agagatctaa	atacagtggc	ttaaatataa	1200
cagaagcata	tttttctctc	atgtaacagc	tagaggaatg	tgtgtagtcc	agagcatata	1260
aaccacaagt	cattcatgac					1280

<210> 81

<211> 974

<212> DNA

<213> Homo sapiens

<400> 81

ggcagagcc	acaactacca	gaactggagg	gtgtacgacg	tcttgggtgct	caaaggatcc	60
-----------	------------	------------	------------	-------------	------------	----

cagttatctg	caagggctgc	agatggatcc	ccctgcaatg	tcctcctgtg	ctctgtggtc	120
cccagcagac	gcatggactc	tgtgacctgg	caggaaggga	aggggtcccgt	gaggggcccgt	180
gttcagtcct	tctggggcag	tgaggctgcc	ctgctcttgg	tgtgtcctgg	ggaggggctt	240
tctgagccca	ggagccgaag	accaagaatc	atccgttgcc	tcatgactca	caacaaaggg	300
gtcagtttta	gcctggcagc	ctccatcgat	gcttctcctg	ccctctgtgc	ccttccacgg	360
ctgggacatg	ccttggattc	tgatgctgct	gttcacaatg	ggccagggag	ttgtcatcct	420
ggccttcaga	tcgtgtctgg	aggcagaggt	ccgtgggggt	ccaggcagag	gaaaccgaag	480
cgggtgttaa	actgtgttgg	aagccccagc	agttttttga	aagaggccgt	gaccacctgt	540
ggcgagggca	gacccccagc	aggcctggaa	cagatcaagc	tacctggaaa	ccccccagtg	600
accttgattc	accaacatcc	agcctgcgtc	gcagcccatc	attgcaatca	agtgagagaca	660
gagtcgggtg	gagacgtgac	ttatccagcc	cacagggatg	tacctgggag	acctgtgcaa	720
cagcgccgtg	gcaagccatg	tggccccctg	aggcattttg	gctgcagcag	ctaccgccct	780
gacctgtttt	ttgccaggac	tgtggagcgg	atagggggag	taggagtaga	gaaggggaaca	840
agggagcaag	ggaacaaggg	acatttgaac	atttaatgtg	agaagacaaa	catcctttgt	900
gagtcattaa	aatttatgaa	ccacttaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	960
aaaaaaaaaa	aaaa					974

<210> 82

<211> 1955

<212> DNA

<213> Homo sapiens

<400> 82

ggcagtgctc	ccaaggcacc	gaaaccgagg	cgggggtctc	ggtccctccg	cgcaaggagg	60
gagggcgacc	gtacgtggca	ggactcaccg	ccccgcacgt	ggcaggactc	accgccccgc	120
gccgtgttct	ccgagccatg	gcgccagcgc	tgtggcgggc	ctgcaacgga	ctcatggccg	180
ccttctctgc	gctggcggcc	ttgggtgcagg	tggtgtacac	aatccctgca	gtactgaccc	240
tgcttgttgg	acttaaccct	gaagtcacag	gtaatgttat	ttggaaaagt	atctctgcaa	300
tacacatact	cttttgtacg	gtgtgggctg	ttggccttgg	gtccctacctc	ttgcatcgta	360
cacaacagaa	catcttaccat	gaggaagaag	gcaggagctg	tctgggtctgg	tgattattac	420
agcatggatt	atcctgtgcc	acagttcctc	aaagaatcca	gttgggtggaa	gaattcaatt	480
ggctattgcc	attgtaatca	cacttttccc	atttatctca	tgggtctaca	tatatattaa	540
caaggaaatg	cggctcctct	ggccaactca	ctgcaagaca	gtaattttaa	ttaattcaag	600
aacttcgttt	ttaaaatgaa	tatttttcaat	caatttttta	ttaacattag	gggaacaagc	660
caggaattta	tttcaggtaa	tttgggctaa	tagtttttaa	actccaaata	acttttttaag	720
ggtgcataata	attcgatgta	agattggatg	ggacaagtaa	gaaatgggtc	gataattttcc	780
agacactttc	tgcaagggtc	tgtgtcataa	tgtagtggaa	aaggctagaa	ataaagttta	840
aaaatacagt	tctaacttaa	ctttgtacta	tgttatttgg	gcaatatata	aacctcctgg	900
tggatattat	ctataaaaata	ggataatgcc	agatctactt	acttacacag	taacaaggat	960
caatctagat	aatgtaagaa	cactctgaag	atataaagtg	tttggaaaaga	ttacggaggg	1020
ctgcccataga	ttaaaataga	ggggagcagg	gattggtaat	atactgaaat	agacattcaa	1080
gagagggtat	accccgatct	tttttttttt	ttttaagaca	gtcccactct	attgcccagg	1140
ctggagtgcg	gtggcatgat	catggctcac	tgcagtcctg	acctcccggg	gctcagggtga	1200
tcctcctacc	tcagcctccc	aagaagctac	gactacaagt	gtgcaccacc	atgcccggct	1260
aaattttttt	aaattttttg	tagaggcagg	gtttcacctg	gttgtccagg	ctgggtcttga	1320
attcctgagt	tcaagcgatc	tgtccacctt	ggcctcccaa	agtgtctggg	ttacaggtgt	1380
gagccaccat	gcccagtcga	gccctaaact	aatcttatcc	agagctggca	tagtgcatgt	1440
aactaaagga	gaactctaga	tgaatcaaat	gagcaggcat	gtcttggaag	gaaaggggaag	1500
ctggatagaa	taaagggaatt	agggaacaaa	caagaaggca	atagggacta	taactacatt	1560
ctaaatgaga	aataagggtca	aaatctatat	acattcttta	taaattggatg	tccaaagtaa	1620
tcctagggag	gagagctttt	tttttttcat	ttccattttc	atttaaaaat	ggatacttga	1680
ttattggaaa	acttacaatt	gtgtttggaa	caacttgggt	atttgaaatc	aattttccaa	1740
ttgcaaattt	tatgatacct	aaatacacag	aaagtatttc	caatgaaaat	ttagtgtcca	1800
aatgagtggg	gctataaatg	taaaatacac	tggattttta	agacatggta	gaaaaagaac	1860
agaaaatatt	ttttgtattg	attacatgtt	gaaacaatat	tttggctata	atgggttaaa	1920
tatactatta	aagttgaaaa	aaaaaaaaaa	aaaaa			1955

<210> 83
 <211> 638
 <212> DNA
 <213> Homo sapiens

<400> 83
 ggcacgagag aggtcctggg gtaagagaaa aaaagtagtt atagcacttc gtccagcact 60
 gacagcagcc gaacaaatgc tgaaaatctg actgtgtgac agaacgtatc actgatgact 120
 gatagaaaagc cctctttcac tctgattacc cactcactac atgaagtcct gaaaataaca 180
 gagaaactgt tatatctttt taatgattta ttgcaagta ttgagatttg acctgaaaaa 240
 caatgaaaca catgaacaca cttccgattt tctcctcgct gattagcttc ctgctgctg 300
 tcagtgtctg acgaagtgtc ataactactt tatgtaacat tacagaacag cttagaggtcc 360
 tggggtaaga gaaaaaaagc acatcacaaac aaatgtgaaa gccttcatta ttacacgttc 420
 cagtttctct cgctgtgtag gcataagcta atggtttatt ttcagaaagc tgctgaaac 480
 gttgtcttgt attcttctag gaagaacttt aattcctcct gaggaactct actttctgag 540
 ccaaactgct aattttctgc ggaactgtct agaagatcat tcaagagacc ctgcagttgc 600
 actttctcgt aaaagttaaa aaaaaaaaaa aaaaaaaa 638

<210> 84
 <211> 859
 <212> DNA
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (27)
 <223> n equals a,t,g, or c

<400> 84
 ccgggtcgac ccacgcgtcc ggcagangcg ggactgtcgt ctggggggagc cgcccaggag 60
 gtccttcagg ccgacccag accctggctg gccaggatga agtatctccg gcaccggcgg 120
 cccaatgcc cctcattct ggccatcgcc gctttcaccc tctcctctt cagtctgcta 180
 gtgtcaccac ccactgcaa ggtccaggag cagccaccgg cgatccccga ggccctggcc 240
 tggccactc caccacccg ccagccccg gccccgtgcc atgccaacac ctctatggtc 300
 acccacccg acttcgccac gcagccgcag caggttcaga acttcctcct gtacagacac 360
 tgccgccact ttccccgtct gcaggacgtg cccccctcta agtgcgcgca gccggtcttc 420
 ctgctgctgg tgatcaagtc ctccccatgc aactatgtgc gccgcgagct gctgcccgcg 480
 acgtggggcc gcgagcgcaa ggtacggggg ttgcagctgc gcctcctctt cctgggtggc 540
 acagcctcca acccgcacga ggcccgcaag gtcaaccggc tgcctggagct ggaggcacag 600
 actcacggag acatcctgca gtgggacttc cagactcct tcttcaacct cagcgtcaag 660
 caggtgcgct ggactggggc cacctgatcg gggccacct tcttcttctg ccaaattacc 720
 actccactcc agcctgggca acaaaagcga aaactccatc tccaaaaaaa taataataat 780
 aataataaaa taaaaatcac acaaaggcca aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 840
 aaaaaaaaaa aaaaaaaaaa 859

<210> 85
 <211> 1129
 <212> DNA
 <213> Homo sapiens

<400> 85
 gctgcttccc aaggaccatg aaactcctgc tgctggctct tctatgctt gtgctcctac 60
 cccaagtgat ccagcctat agtgggtgaaa aaaaatgctg gaacagatca gggcactgca 120

ggaaacaatg	caaagatgga	gaagcagtga	aagatacatg	caaaaatctt	cgagcttgct	180
gcattccatc	caatgaagac	cacagggcgag	ttcctgcgac	atctcccaca	cccttgagtg	240
actcaacacc	aggaattatt	gatgatattt	taacagtaag	gttcacgaca	gactactttg	300
aagtaagcag	caagaaagat	atgggttgaag	agtctgaggc	gggaagggga	actgagacct	360
ctcttccaaa	tggtcaccat	agctcatgac	ttcctctcgg	ctatcactca	ccccgtcct	420
cagagtgcata	aactaagtca	catacagata	aagcactgaa	aacaccacag	tgaccctccc	480
acccccacc	aatatgtaat	tctattaata	gaaacagctg	tgtaaagaag	tctaaaattt	540
tcactatttc	caatgataaa	ctcttcagt	ctcttcttga	aatgtcacat	tatttcccaa	600
caagttatac	ctatttttag	tattcttggt	gctagtgcga	tgacaaactt	caatagctag	660
ttgctattcc	aacaacaatt	tcttcatgta	tcgttctgtc	ttctcaacag	ctgtctttca	720
tggcagcata	agtggtcag	atcaaaaattc	taaatcttgc	atctgtgaga	gtagctacta	780
tgacactaaa	agcttttttt	ctagaacagg	agacacttca	ggtgaagcat	tcattctcct	840
actaactatg	gccttggagc	cagggttttat	ctctcactgt	aggaaaattg	ccgccccagg	900
tgtgagctat	gaagactcct	ttttgccccca	gtggcttttg	ggttgaaatg	ctgtcgaaaa	960
gcttttatgg	ctctgtagac	ccatcttttt	gaccaagcct	tgatcacaca	tggacatcca	1020
agggtaatca	tggacccccca	attgtgggtg	aaaggatgga	tcatttatct	acctgattac	1080
tgagagcttt	atttgtctcc	ctctgatagc	aaaaaaaaaa	aaaaaaaaaa		1129

<210> 86

<211> 2674

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (2607)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (2611)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (2621)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (2634)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (2650)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (2660)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (2669)

<223> n equals a,t,g, or c

<400> 86

gatccctccc	atctcacagt	acctcacagg	tctcttcccc	cgagcagtg	attgctggag	60
cgaggagaag	ctcacgaatc	agctgcaggt	ctctgttttg	aaaaagcaga	gatacagagg	120
cagaggaaaa	gggtggactc	ctatgtgacc	tgctcttaga	gcaagacaat	caccatctga	180
attccagaag	ccctgttcat	gggtggggat	atcttctcga	ctgcatggaa	tcagaaagaa	240
gcaaaaggat	gggaaatgcc	tgcatctccc	tgaaaagaat	tgcttatttc	ctatgtctct	300
tatctgcgct	tttgctgact	gaggggaaga	aaccagcgaa	gccaaaatgc	cctgccgtgt	360
gtacttgtac	caaagataat	gctttatgtg	agaatgccag	atccattcca	cgcaccgttc	420
ctcctgatgt	tatctcatta	tcctttgtga	gatctgggtt	tactgaaatc	tcagaaggga	480
gttttttatt	cacgccatcg	ctgcagctct	tgctttcac	atcgaaactcc	tttgatgtga	540
tcagtgatga	tgcttttatt	ggcttccac	atctagagta	tttattcata	gaaaacaaca	600
acatcaagtc	aatttcaaga	catactttcc	ggggactaaa	gkcatatt	cacttgagcc	660
ttgcaaaaca	caatctccag	acactcccaa	aagatatttt	caaaggcctg	gattctttaa	720
caaatgtgga	cctgaggggt	aattcattta	attgtgactg	taaactgaaa	tggtagtgg	780
aatggcttgg	scacaccaat	gcaactgttg	aagacatcta	ctgcgaaggc	cccccagaat	840
acaagaagcg	caaaatcaat	agtctctcct	cgaaggattt	cgattgcata	attacagaat	900
ttgcaaaagtc	tcaagacctg	ccttatcaat	cattgtccat	agacactttt	tcttatttga	960
atgatgagta	tgtagtcata	gctcagcctt	ttactggaaa	atgcattttc	cttgaatggg	1020
accatgtgga	aaagaccttc	cggaattatg	acaacattac	aggcacatcc	actgtagtat	1080
gcaagcctat	agtcattgaa	actcagctct	atgttattgt	ggcccagctg	tttggtggct	1140
ctcacatcta	taagcgagac	agttttgcaa	ataaattcat	aaaaatccag	gatattgaaa	1200
ttctcaaaat	ccgaaaaccc	aatgacattg	aaacattcaa	gattgaaaac	aactgggtact	1260
ttgttgttgc	tgacagttca	aaagctgggt	ttactaccat	ttacaaatgg	aacggaaacg	1320
gattctactc	ccatcaatcc	ttacacgcgt	ggtacaggga	cactgatgtg	gaatatctag	1380
aaatagtcag	aacacctcag	acactcagaa	cgccctcatt	aattctgtct	agtagttccc	1440
ascgtcctgt	aatttatcag	tggaacaaag	caacacaatt	attcactaac	caaactgaca	1500
ttcctaacat	ggaggatgtg	tacgcagtga	agcatttctc	agtgaagggg	gacgtgtaca	1560
tttgcttgac	aagattcatt	ggtgattcca	aagtcattgaa	atggggaggc	tcctcgttcc	1620
aggatattca	gaggatgcca	tcgcgaggat	ccatgggtgt	ccagcctctt	caaataaata	1680
attaccaata	tgcaattctt	ggaagtgtat	actcctttac	tcaagtgtat	aactgggatg	1740
cagagaaagc	caaatttgtg	aaatttcagg	aattaaatgt	tcaggcacca	agatcattca	1800
cacatgtgtc	cattaataag	cgtaattttc	tttttgcttc	cagtttttaag	ggaaatacac	1860
agattttaca	acatgtcata	gttgacttaa	gcgcattgaga	caccaaattc	tgtggctgcc	1920
atcagaaatt	ttctacagta	catgaccccg	atgaactcaa	tgcatgatga	ctcttcttat	1980
cacacttgca	aatgaatgcc	tttcaaacat	tgagactgct	agaaccaagc	actaccagta	2040
tctccatcct	taactgtcca	gtccagtgtg	gtgggaagtt	accttttata	agacaaaatt	2100
taattgtgta	actgttcttt	gcagtgaaga	tgtgtaaata	agcgtttaat	ggtatctgtt	2160
actccaaaaa	gaaatattaa	tatgtacttt	tccattttat	tattcatgtg	tacagaaaca	2220
actgccaat	aaaaatgtta	catttttctt	cataaaaaaa	aaaaaaaaaa	aactcgaggg	2280
ggggcccggt	acccaattcg	ccctatagtg	agtcgtatta	caattcactg	gccgtcggtt	2340
tacaacgtcg	tgactgggaa	aaccctggcg	ttacccaact	taatcgccct	gcagcacatc	2400
cccctttcgc	cagctggcgt	aatagcgaa	aggccgcacc	gatcgccctt	cccaacagtt	2460
gcgcagcctg	aatggcgaa	ggcaaattgt	aagcgttaat	attttgttaa	aattccgcgt	2520
taaattttgt	taaatcagct	cattttttaa	cccaataggc	cgaatttcgg	caaaaatccc	2580
ttattaatca	aaagaaatag	aaccganaat	nggggttgaa	ntgttggttc	caantttggg	2640
aaacaaaaan	tcccacttan	tttaaaagna	aacg			2674

<210> 87

<211> 1636

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (1624)

<223> n equals a,t,g, or c

<400> 87

tcgaccacg	cgtccggctg	agtgtgagct	gagcctgccc	caccaccaag	atgatcctga	60
gcttgctgtt	cagccttggg	ggccccctgg	gctgggggct	gctgggggca	tgggcccagg	120
cttccagtac	tagcctctct	gatctgcaga	gctccaggac	acctgggggtc	tgggaaggcag	180
aggctgagga	caccagcaag	gaccccggtg	gacgtaactg	gtgcccctac	ccaatgtcca	240
agctggtcac	cttactagct	ctttgcaaaa	cagagaaaatt	cctcatccac	tcgcagcagc	300
cgtgtccgca	gggagctcca	gactgccaga	aagtcaaagt	catgtaccgc	atggcccaca	360
agccagtgtg	ccagggtcaag	cagaagggtg	tgacctcttt	ggcctggagg	tgctgcccctg	420
gctacacggg	ccccaaactgc	gagcaccacg	attccatggc	aatccctgag	cctgcagatc	480
ctggtgacag	ccaccaggaa	cctcaggatg	gaccagtcag	cttcaaacct	ggccaccttg	540
ctgcagtgat	caatgaggtt	gaggtgcaac	aggaacagca	ggaacatctg	ctggggagatc	600
tccagaatga	tgtgcaccgg	gtggcagaca	gcctgccagg	cctgtggaaa	gccctgcctg	660
gtaacctcac	agctgcagtg	atggaagcaa	atcaaacagg	gcacgaattc	cctgatagat	720
ccttgaggca	ggtgctgcta	ccccacgtgg	acaccttctt	acaagtgcac	ttcagcccca	780
tctggaggag	ctttaaccaa	agcctgcaca	gccttaccga	ggccataaga	aacctgtctc	840
ttgacgtgga	ggccaaccgc	caggccatct	ccagagtcca	ggacagtgcc	gtggccagggt	900
ctgacttcca	ggagcttggg	gccaaaattg	aggccaaggt	ccaggagaac	actcagagag	960
tgggtcagct	gcgacaggac	gtggaggaac	gcctgcacgc	ccagcacttt	accctgcacc	1020
gctcgatctc	agagctccaa	gccgatgtgg	acaccaaatt	gaagaggctg	cacaaggctc	1080
akgaggcccc	agggaccaat	ggcagtctgg	tggtggcaac	gcctggggct	ggggcaaggc	1140
ctgagccgga	cagcctgcag	gccaggctgg	gccagctgca	gaggaacctc	tcagagctgc	1200
acatgaccac	ggcccgcagg	gaggaggagt	tgagctacac	cctggaggac	atgaggggcca	1260
ccctgaccgg	gcacgtggat	gagatcaagg	aactgymctc	cgaatcggac	gagactttctg	1320
atcagattag	caagktgkwg	cggcagggtg	aggagctgca	ggtgaaccac	acggcgctcc	1380
gtgagctcgc	cgtgatcctg	atggagaagt	ctctgatcat	ggaggagaac	aaggaggagg	1440
tggagcggca	gctcctggag	ctcaacctca	cgtgcagca	cctgcagggt	ggcatgccga	1500
cctcatcaag	tacgtgaagg	actgcaattg	ccagaagctc	tatttagacc	tggacgtcat	1560
ccgggagggc	agagggacgc	cacgcgtgcc	ctggaggaga	cccagggtgag	cctggacgar	1620
cgngggcaag	ctggac					1636

<210> 88

<211> 1639

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (12)

<223> n equals a,t,g, or c

<400> 88

gtgacactat	anaagtacgc	ctggcagggt	accggtccgg	caattcgcg	ccgcgtcgac	60
gtcaggcggg	ccgtgggttc	ccggaggggg	tcttgaggcg	ccatttcaag	tcgccccag	120
cctctcccac	agcactcctg	ttttcccggg	cctcatcatg	gccacggcc	ctcagtcgct	180
gtggagcctg	ggcttcacag	tgacactcac	gtttgaactc	ccggctcggt	gtgtgcttgg	240
tagaatttgt	catccaatac	aggcgtgtaa	cacgggcttg	atgacacca	ccccacaggg	300
cccctgcagg	accgagatga	tgtccaatga	caagccctgg	cttcagcca	atgtcctgc	360
ccacatctct	ctcccaggag	ccaggcttac	ctctacctgt	gcacctgggc	tgtgactcat	420
gactggaatg	atctggctgg	gcctgttccc	ccaccagact	tattttcagg	cgccccagca	480
gccaccaaac	atctgtcaac	tgaagtaatg	aacctgcagt	tgagaggcag	ctaaacctag	540
gttgaagggt	aggagagacg	ctgagttgag	gtcaaatccc	cccggccagt	tagcctttct	600
gagcctatct	cctcaattgt	aaaaggaaga	caatcatgat	gctgacctca	gaatcsagcg	660
aggaggaagt	gaggaggtgc	atgtgaagca	ttgtgcgtga	ctggtggggc	taactgggct	720

ctggaagcgg	tartctctggg	gccctaaccc	ctttctgcct	catgctaatt	gacctatggc	780
atgtccagt	acatcatgac	tgaaaaaatg	atttgaaart	ataatcgctg	tcaggaattg	840
tctcctggc	tcaaactcct	aggcttaagg	aatttgccca	ccttagccta	ccaaagggcg	900
gggcttaca	gcatgarcca	tggcaccg	ccccaaagt	tttttctat	tctctcagtc	960
macagttaca	cagaaaattt	ctgtgaccac	tggtcacgaa	agggagtggg	ggctctctccc	1020
taccggcaac	caagcagtcg	attctgcagt	ggacaccagc	tgggtgttct	cttattgaat	1080
taattctgac	actatctgtc	cagagatagc	attagattcc	acagggtgag	gacttagtcc	1140
ccacttgctc	cttatttctg	atgctgatca	caagamctar	gttatcttgc	cggtgattct	1200
gactgactgg	ctattaatta	gggtttctgt	gacccactcc	ttgggttcaa	ttaatttgct	1260
agagaactcc	ctcatggaat	tcagagaaac	acattttacca	gcttattata	aaggctgcta	1320
caaaggatac	agatgagatg	cgcagagcaa	cgtgtgggat	ggagtgcaga	gcttccgcgc	1380
cctctcctgg	agcaccactc	ttcaggaacc	tccatgtgtt	cagctattca	gaagctccct	1440
ggaccacgtc	ctttcgggtt	tttatggaag	cttcattatg	tagacatgat	taattatacc	1500
attggtcatt	ggtgatcaac	ttaaccttca	gcccttctcc	cctcccgagg	gttggagggt	1560
ggggctgaaa	cattccaact	tacaggcccg	tcgacgcggc	cgcgaattcc	cgggtcgacg	1620
agctcactag	tcggcggcc					1639

<210> 89

<211> 1860

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (1846)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (1848)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (1853)

<223> n equals a,t,g, or c

<400> 89

ctcaccagms	ggaaagtacg	agtcggctca	gcctggagg	acccaaccag	agcctggcct	60
gggagccagk	atggccatcc	acaaagcctt	ggtgatgtgc	ctgggactgc	ctctyttcct	120
gttcccagg	gcctgggccc	agggccatgt	cccaccggc	tgcagccaag	gcctcaaccc	180
cctgtactac	aacctgtgtg	accgctctgg	ggcgtggggc	atcgtcctgg	aggcctgggc	240
tggggcgggc	attgtcacca	cgtttgtgct	caccatcatc	ctgggtggcca	gcctcccctt	300
tgtgcaggac	accaagaaac	ggagcctgct	ggggaccag	gtattcttcc	ttctggggac	360
cctgggcctc	ttctgcctcg	tgtttgctg	tgtggtgaag	ccgacttct	ccacctgtgc	420
ctctcggcgc	ttctcttttg	gggttctgtt	cgccatctgc	ttctcttgct	tggcggtcca	480
cgtctttgcc	ctcaacttcc	tggcccggaa	gaaccacggg	ccccggggct	gggtgatctt	540
cactgtggct	ctgctgctga	ccctggtaga	ggtcatcatc	aatacagagt	ggctgatcat	600
cacctgtgtt	cggggcagtg	gcgagggcgg	ccctcagggc	aacagcagcg	caggctgggc	660
cgtggcctcc	ccctgtgcca	tcgccaacat	ggactttgtc	atggcactca	tctacgtcat	720
gctgctgctg	ctgggtgcct	tcctgggggc	ctggcccgcc	ctgtgtggcc	gctacaagcg	780
ctggcgtaag	catggggctc	ttgtgctcct	caccacagcc	acctccgttg	ccatatgggt	840
ggtgtggatc	gtcatgtata	cttacggcaa	caagcagcac	aacagtccca	cctgggatga	900
ccccacgctg	gccatcgccc	tcgcccga	tgcttggg	ttcgtcctct	tctacgtcat	960
ccccgaggtc	tcccaggtga	ccaagtccag	cccagagcaa	agctaccagg	gggacatgta	1020
ccccaccg	ggcgtgggct	atgagaccat	cctgaaagag	cagaagggtc	agagcatggt	1080

cgtgggagAAC	aaggccctttt	ccatggatga	gccgggttga	gctaagagggc	cggtgtcacc	1140
atacagcggg	tacaatgggc	agctgctgac	cagtgtgtac	cagcccactg	agatggccct	1200
gatgcacaaa	gttccgtccg	aagagcttac	gacatcatcc	tcccacgggc	caccgccaac	1260
agccaggtga	tgggcagtg	caactcgacc	ctgcgggctg	aagacatgta	ctcggccccag	1320
agccaccagg	cgccacacac	gccgaaagac	ggcaagaact	ctcaggtctt	tagaaacccc	1380
tacgtgtggg	actgagtcag	cggtggcgag	gagaggcgg	cggtattggg	gagggccctg	1440
aggacctggc	cccgggcaag	ggactctcca	ggctcctcct	ccccctggca	ggcccagcaa	1500
catgtgcccc	agatgtggaa	gggcctccct	ctctgccagt	gtttgggtgg	gtgtcatggg	1560
tgtccccacc	cactcctcag	tggttgtgga	gtcaggagac	caacccagc	ctcctgccag	1620
gatcacctcg	cgggtcacac	tccagccaaa	tagtgttctc	ggggtgggtg	ctgggcagcg	1680
cctatgtttc	tctggagatt	cctgcaacct	caagagactt	cccaggcgct	caggccctgga	1740
tcttgtcctc	ctgtgaggaa	caagggtgcc	taataaatac	atttctgctt	tattaaaaaa	1800
aaaaaaaaaa	aaaaaaactc	gagggggggc	ccgtacccaa	tcgccngnga	tgntagtata	1860

<210> 90

<211> 839

<212> DNA

<213> Homo sapiens

<400> 90

ggcacgaggg	ctacgatcct	acagtggaga	atagatgagt	acagcattct	gccctattca	60
ttcatcattg	gggtccatgg	ttatgtgctt	gtgtattctg	tcacctctct	gcatagcttc	120
caagtcattg	agagtctgta	ccaaaagcta	catggaaggc	catgggaaaa	cccgggtgcc	180
agtgtttcta	gtggggaaca	aggcagatct	ctctccagag	agagaggtac	aggcagttga	240
aggaagaag	ctggcagagt	cctgggggtg	gacatttatg	gagtcattct	ctcgagagaa	300
tcagctgact	caaggcatct	tcaccaaagt	catccaggag	attgcccgtg	tgggagaatt	360
cctatgggca	agagcgtcgc	tgccatctca	tgtgagccct	tgggtgtggg	gtaactgcct	420
tgcttctgcc	cccggcactt	gccatgttcc	agtggggggc	agatcctcag	gacttcacgg	480
gtatggttgc	cagctgtgtt	cctggccctc	ggacacacag	tgtggcatcc	tcattgtttg	540
acactttccc	caggctccag	tggcctggat	gtcaatgttt	acaaaggggc	aaggacctct	600
catggacact	ggcctctagc	cctctgtttt	tgtttgatga	attctgttat	aacctatggg	660
gtcaggatat	gagtcctggg	cattatttat	ccaggaccca	tcctcttggg	tgggttttgg	720
gtgttggctg	ggtaagggga	gccggggact	tctgaaatag	agctggctcc	ctggggtgac	780
aatgtatata	tgcaataaaa	ttgagaaatc	ttttgtgtgt	gaaaaaaaaa	aaaaaaaaaa	839

<210> 91

<211> 1145

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (386)

<223> n equals a,t,g, or c

<400> 91

aattcggcac	gaggacatat	tggccattta	ctctactaat	aaaagagtac	tatctactca	60
gtgtcattta	ctgttactgt	agtaattaat	gccttttgaa	gaatcttttg	aaatagttct	120
caaattggta	ccactacttg	gtttggaatt	attttttttt	cttttcataa	tcaatggtta	180
tataaatgta	tattgtccta	gtcagtattt	tatatatgct	aaggactcac	tagctggctt	240
ggcactaata	cctcaataaa	aggaataact	cttttggaat	catgaaacaa	aagtgartaa	300
acctccaagt	tatttttcca	accaaccttc	tttgaaaaat	cttggtgag	tactcaaat	360
caagacatgt	tataaaatta	tctgtnatat	tggtagaaca	tataattgt	yctaataata	420
atttycaaat	attcagtgka	acygtaagka	tgagaatata	ggttgaatat	cycttatcca	480
aatgtcttgg	gaccagaagt	cttttggatt	ycaaattttt	aaatattttac	atcatactta	540

ccagtttaac	atccctaatt	caaaaattca	aaattcagaa	tgctccaata	atcgtttcct	600
ttgascatca	tgtcagtgtc	caaaaagtgt	cagattttga	ggtatttcag	atctttgaat	660
taggaagact	caacttgtac	tatcattcta	tagactttat	gattgggtag	actacatgag	720
tattgaaccc	agaaatcatt	gtctagcaaa	agccagtata	gtgattaatt	accctgtgac	780
tattatataa	tgttcaaaaa	agctaacata	ttagaatgtc	cttagcgtgc	agagagcaaa	840
cagagacaaa	aagaaaagtt	accctgaaaa	gtttgtcaga	aaaatagaat	atcagacgct	900
raactactca	tccagaattt	tgtrcaaaaa	gaaaaataag	ataaaattca	ctggtagaca	960
aaaagtagta	acataccagt	ttgtaatttc	tcagtttcaa	accatgaata	tgtatttgta	1020
tacaaaaaat	catttcagga	gtcagagaag	gaggatatgc	cttttatgtg	gagactttta	1080
acataaaatt	ggaaaaaaa	aaaaaaaaa	actcgtaggg	gggggcccg	acccaatcgt	1140
cctgt						1145

<210> 92

<211> 2050

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (515)

<223> n equals a,t,g, or c

<400> 92

aagaatggca	taaattcatc	cagcttctta	cagaattcca	aatgcggaat	gtagattttt	60
tatatagtaa	tcttgagttt	attctaccat	taccagttga	taccattcca	gaaactaaaa	120
acttttggg	cccatcagta	actgtggatg	ccagtgcagc	aacaaaaagt	atgaattgtc	180
ttgctaggaa	acactctgaa	agagaacagc	cattgaaaaa	gtcccagaaa	aagaaacaaa	240
agaaaacatt	ggtaatat	gatgatagtg	atctatttga	cactgacttg	gactttcctg	300
atcaatctat	tagcctgtcc	tctgtatcat	cttccctcaa	tgacagaaga	agcaaaaccg	360
gagacgaaga	aagcaaaagc	agagacaaa	gaaacaatcc	agagacaaa	aaatctattc	420
cttgtcctcc	taaaacaact	ggcaggaaaa	aaatgttctg	ccctgtttct	cattgtttta	480
attctctctc	tgagttcatg	ggataacatg	tcctncttag	atgcactttt	aactgatgta	540
aggggracaaa	acaaatacgg	tagaaatgac	tttagttgga	caaatggaaa	ggttacaagt	600
ggactttgtg	atgagtttag	tcttgagagt	aatgatggat	ggacttctca	aagctctgga	660
gaattaaagg	cagctgcaga	agctctcagc	tttactaaat	gttcttctgc	tatttcaaaa	720
gcatkggaaa	ccttgaaatc	ttgcaagaaa	ttaggaagag	atccaaccaa	cgatcttact	780
ttttatgttt	tcacaaaagc	caataatgta	tacttttagt	agtcagcagc	taatttagac	840
aatgcttggg	agaggatata	agtcattaaa	agtgtatttt	cgagtgcagc	tcttctctat	900
gtgggtaata	gacaagctag	tataattgaa	tacctgcaa	cccttcgaaa	catctgtaag	960
actgagaagc	taaaagaaca	aggaaaaagt	aaaagaagat	tcctgcacta	tttgaagga	1020
attcatcttg	acattccaaa	agagactgtg	aatacttttg	cagctgactt	cccttaattg	1080
tccatactaa	caatgctttg	tatagattat	catgtggtcc	ttaagataca	tttttatatt	1140
atgtggatct	tcattggaaa	gtatatattc	cgatgtacat	tttaaacaaa	caatttgtat	1200
atttttttat	tgccgggtaa	atattttaa	tatttgagtt	acaaatttta	tatatgattg	1260
taattttttt	tctgaatttt	ttgtattatc	tgatttagct	ttgttgaggt	attttttgta	1320
tgtgagtgaa	ctgtttctgg	aaggtagagt	tcattaagat	gaactcccta	tttcaagtgt	1380
ttatattata	tattagctta	atattcagat	acattatctt	ggctgctaac	attagtgtca	1440
ctaaagtgtg	tatacaatct	cccactgcta	aatttgactg	gctttacaaa	aacaaaaaca	1500
ttatctgggt	aatttatatt	ttaacctaaa	agttaaggat	cctcatattg	tacagttttt	1560
tttgtgtgct	tttttttttt	tttttgagac	ggaatcttgc	tctgtcaccc	aggctggagt	1620
gcagtggcct	ggtatcggct	cagtgcgaact	tttgccctcc	gggttcaagc	gattatcctg	1680
cctcagcctc	ctgaatagck	gggattacag	gcattgtgcca	ccttgcccag	ctaatttttg	1740
tatttttaga	agagacaggg	ttttaccatg	ttgggttaggc	tggtctctta	actcckgacc	1800
tcaagtgtac	catctgcctc	ggcctcccaa	agtgtctgga	tcacaggcgt	gagccacctc	1860
acctggccta	tattgtacag	ttttgaacag	tatagatgca	tacctgttta	caaattgtgt	1920
tgaagataga	tattttttacc	tcttatttgt	tcaatttact	ttktcttgta	ttaattagta	1980

tattgatcta attaaagggt aaagctaaag gctttatgaa atgtttaaaa aaaaaaaaaa 2040
 aaaactcgag 2050

<210> 93
 <211> 1173
 <212> DNA
 <213> Homo sapiens

<400> 93
 tcgacccacg cgtccgaaac aaaggaaaat atccccaag ttgttttcta gatttgtggc 60
 ttttaagaaaa acaaaacaaa acaaacacat tgttttttctc agaaccagga ttctctgaga 120
 ggtcagagca tctcgctgtt tttttgtgtg tgtttttaaaa tattatgatt tggctacaga 180
 ccaggcaggg aaagagaccc ggtaattgga ggttgagcct cggggggggg gcaggacgcc 240
 ccggtttcgg cacagcccgg tcaactcacgg cctcgctctc gcctcacccc ggctcctggg 300
 ctttgatggg ctggtgccag tgcctgtgcc cactctgtgc ctgctgggag gaggcccagg 360
 ctctctggtg gccgccctg tgcacctggc cagggggaagc ccgggggtct ggggcctccc 420
 tccgtctcgg cccacctttg cagaataaac tctctcctgg ggtttgtcta tctttgtttc 480
 tctcacctga gagaacgca ggtgttccag aggttctctt gcagacaaag cacccttgca 540
 cctcctatgg ctacaggatga gggatgcccc caggcccttc tggttggtag tgagtgtgga 600
 cagcttccca gctcttcggg tacaacctg agcaggtcgg gggacacagg gccgaggcag 660
 gccttcgggg cccctttcgc ctgcttcggg gcaggagca ggcctgggtg cctcgctcca 720
 cccacccacg ctgctgtcac ctgaggggaa tctgcttctt aggagtgggt tgagctgata 780
 gagaaaaaac ggccttcagc ccaggctggg aagcgccttc tccaggtgcc tctccctcac 840
 cagctctgca cccctctggg gagccttccc cacttagct gtctcctgcc ccaggagggg 900
 atggaggaga taatttgctt atattaaaaa caaaaaatgg ctgaggcagg agtttgggac 960
 cagcctgggc tatatagcaa gaccccatca ctacaaattt tttacaaatt agctaggtgt 1020
 ggtggtgcgc acctgtggtc ccagctactc gggaggctgt ggtgggagga ttgcttgagt 1080
 ccaggagggt gaggtgcag tcagctcaga ttgcaccact gcactccagc ctgggcaaca 1140
 gagcgagacc cgtctctcaa aaaaaaaaaa aaa 1173

<210> 94
 <211> 822
 <212> DNA
 <213> Homo sapiens

<400> 94
 ggcacgaggt tccctctccc cagagccatc ggccaggtac caaagctcag ctgtatggat 60
 tcccaacagg aggacctgcg ctccctggg atgtgggtct cattgtactt tggaatcctg 120
 gggctgtgtt ctgtgataac tggaggggtg attatctttc tgcactggag gaagaacttg 180
 aggcgggaag agcatgcca gcagtgggtg gaggtgatga gagctgccac attcacctac 240
 agccattgt tgtactggat taacaagcga cggcgctacg gcatgaatgc agcatcaac 300
 acggggcctg cccctgctgt caccaagact gagactgagg tccagaatcc agatgttctg 360
 tgggatttgg acatccccga aggcaggagc catgctgacc aagacagcaa cccaaggcg 420
 gaagcccctg ctccctgca acctgcactg cagctggctc cacagcagcc ccaggccaga 480
 tccccattcc cacttcccat ctttcaggag gtgccctttg cccacccctt gtgcaacctc 540
 cccccctgc tgaaccactc tgtctctat cctttggcca cctgtcctga aaggaaatgtt 600
 ctcttccatt cctcctgaa tctggcccag gaagaccata gcttcaatgc caagcctttt 660
 ccttcagaac tgtagcctcc tctcactgaa ggtgggagct gcaggaatca ggtgcagagt 720
 aggaaatgga actaacctca ggaaggtggt attgacagaa gtcaggaccc acctggatgt 780
 catgctatga aacatttgaa gcaaaaaaaaa aaaaaaaaaa aa 822

<210> 95
 <211> 1077
 <212> DNA

<213> Homo sapiens

<400> 95

ggcacgagtt	ggtgggcaat	agcgcttttc	tctcaagggg	cttttgggcta	tgtgctgccc	60
atcatttcat	tcatccttgc	ctggattgag	acgtgggtcc	tggatttcaa	agtgttacct	120
caagaagcag	aagaagaaaa	cagactcctg	atagttcagg	atgcttcaga	gagggcagca	180
cttatacctg	gtggtctttc	tgatggtcag	ttttattccc	ctcctgaatc	cgaagcagga	240
tctgaagaag	ctgaagaaaa	acaggacagt	gagaaaccac	ttttagaact	atgagtacta	300
cttttggttaa	atgtgaaaaa	ccctcacaga	aagtcacoga	ggcaaaaaa	ggcaggcagt	360
ggagtctccc	tgctgcacagt	aaagttgaaa	tgggtgacgtc	cactgctggc	tttattgaac	420
agctaataaaa	gattttattta	ttgtaatacc	tcacagacgt	tgtaccatat	ccatgcacat	480
ttagttgcct	gcctgtggct	ggtaaggtaa	tgtcatgatt	catcctctct	tcagtgaagac	540
tgagcctgat	gtgttaacaa	ataggtgaag	aaagtcttgt	gctgtattcc	taatcaaaag	600
acttaataata	ttgaagtaac	acttttttag	taagcaagat	acctttttat	ttcaattcac	660
agaatggaat	ttttttgttt	catgtctcag	atttattttg	tatttctttt	ctaacactct	720
acatttccct	tgtttttttaa	ctcatgcaca	tgtgctcttt	gtacagtttt	aaaaagtgtta	780
ataaaatctg	acatgtcaat	gtggctagtt	ttatttttct	tgttttgcac	taagtgtatg	840
gcctgaagtg	ttggacttgc	aaaaggggaa	gaaaggaatt	gcgaatacat	gtaaaatgtc	900
accagacatt	tgtattattt	ttatcatgaa	atcatgtttt	tctctgattg	ttctgaaatg	960
ttctaaatac	tcttattttg	aatgccaaaa	tgacttaaac	cattcatatc	atgtttcctt	1020
tgcgttcagc	caatttcaat	taaaatgaac	taaattaaaa	aaaaaaaaaa	aaaaaaa	1077

<210> 96

<211> 2092

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (637)

<223> n equals a,t,g, or c

<400> 96

gaattcggca	yggcgacctt	tgtgagcgag	ctggaggcgg	ccaagaagaa	cttaagcgag	60
gccctggggg	acaacgtgaa	acaatactgg	gctaacctaa	agctgtgggt	caagcagaag	120
atcagcaaa	aggagtgtga	ccttgaagct	catagacttc	tcacacagga	taatgtccat	180
tctcacaatg	atttctctct	ggccattctc	acgcgttggt	agattttggg	ttctacacca	240
gatgggtgctg	gatctttgcc	ttggccaggg	ggttccgcag	caaaacctgg	gaaaacccaa	300
gggaaagaaa	aagctttctt	ctgttcgtca	gaaatttgat	catagattcc	agcctcaaaa	360
tcctctctca	ggagcccagc	aatttgtggc	aaaggatccc	caagatgatg	acgacttgaa	420
actttgttcc	cacacaatga	tgcttcccac	tcgaggccag	cttgaaggga	gaatgatagt	480
gactgcttat	gagcatgggc	tggacaatgt	caccgaggag	gctgtttcag	ctgttgtcta	540
tgctgtggag	aatcacctta	aagatatact	gacgtcagtt	gtgtcaagaa	ggaaagctta	600
tcggttacga	gatggtcatt	ttaaatatgc	ctttggnagt	aacgtgaccc	cgcagccata	660
cctgaagaat	agtgtagtag	cttacaacaa	cttaatagaa	agccctccag	cttttactgc	720
tcctgtgct	ggtcagaatc	cagcttctca	cccacccct	gatgatgctg	agcagcaggc	780
tgcactcctg	ctggcatgct	ccggagacac	tctacctgca	tctttgcctc	cgggtgaacat	840
gtacgatctt	tttgaagctt	tgcaggtgca	caggggaagtc	atccctacac	atactgtcta	900
tgctcttaac	attgaaagga	tcatcacgaa	actctggcat	ccaaatcatg	aagagctgca	960
gcaagacaaa	gttcaccgcc	agcgcttggc	agccaaggag	gggcttttgc	tgtgctaaat	1020
taggatttga	gggtgtggga	ccctaccra	attcattgat	tactgaaaat	tgaatgtttt	1080
ttgggtccac	atttcaaggc	tgaagtgtgt	agtgtatata	taacctttcc	tatggaaatg	1140
tgacattgag	tacattttgt	gttgctgttg	tgaagccatt	aatataaatc	tttggtaatg	1200
acccatatct	ctatatgtat	gtgttcccag	ttgtgggagc	aggcactaat	gaaatcctgt	1260
gcctggaatg	gagatattta	ggtacctgag	gcttagtgtc	ctgtgggtctg	catgtaagat	1320
agatgacatc	ctagaacaaa	gaagctgttt	taacttaatc	cccctgatca	gcaggatatc	1380

tgtgtgttca	gtgacatcat	acattctgtg	tctagaagtc	taaaatttct	gcctttctcc	1440
taaagaatgt	gttcttgc	tttgggtgaa	ataacctaca	cagtgttaaa	aaccagatac	1500
ctcctttagt	gaccagttca	aattttaata	gcgataggta	gcccctgaga	aatttatcac	1560
tataactcca	caggaaatat	gacttgaag	tgctctgtgt	actaaacaaa	ataaagcccc	1620
tctttgcatt	taaaacccaaa	gtcaaaacaa	aactcttgta	atgcaattaa	ttacttyat	1680
gtcttcccat	gactcaagtt	ttgttaaata	tgcccaaaaa	ctttgattgg	cagtttcttc	1740
ggtaattat	tcctatagaa	tgtattttaa	gaaatctata	caaattggat	atatgcttgg	1800
taattctcca	gtttctagga	ggtacctatt	tctaccgttt	caagtgatga	agtgaataa	1860
atttacattc	gatagtgtta	ctgataacaa	acctacttaa	gagatatgtt	gctttttact	1920
taagggatag	tgttgataga	taaattagaa	tgtatagata	ggtttgtgaa	agtctaaata	1980
atgggtgtat	agatatgtat	atatggttca	cayatctgga	tctgtgtatt	tgattttgta	2040
ctttaaatgt	gacaaataaa	ccttttggga	gaaaaaaaac	tc		2092

<210> 97

<211> 1352

<212> DNA

<213> Homo sapiens

<400> 97

ggcacgaggt	gateccacca	cggtggcctc	ccaaagtgtc	gggattacag	gtgtgagccg	60
ctgcccctgg	ccctaaatag	atttttaata	agttttctgg	atgcacacac	tagtaaacac	120
aaatgccaaa	gacttgcctt	ccaattctct	gtctacctct	aactcaggct	gttgccttgg	180
cacagttaaa	caacttttct	agcctcaata	ttttcatctt	caaaatcaaa	aataaaatgt	240
attacatatg	gatctatgac	aaatagtgtat	atattccatg	tgcacgttat	cattaaacat	300
gaattaacca	atattttaaat	cttttttttt	tttttttttt	ttttgagacg	gagcctgggc	360
gacagagaga	gacaccgtct	caaaaaaaaa	aaaaaaaaaa	aaaaaaggta	attatggtaa	420
aatgaggtca	ttaggggtgg	ccctaattct	aattcagtac	aactgatgtc	tttattagaa	480
ttgcaagttt	gggcatcaag	agacacactc	acgcggggaa	gaacacgtga	agacacaggg	540
aggagacagg	gtttacataa	agtcaaggag	aagggcctga	aacattcttc	tttcacagtc	600
tcgcataaag	aaccaaccct	gccgaaacct	tgatctcaga	cttcacgcct	caagaactgt	660
gagaaaataa	atttctgttg	tttaagccat	tcagcttgtg	gaacttttat	atgatgactc	720
tggcaacta	atatggcatt	acacacaggg	gataagaaaa	taatccaaat	aagtatatgt	780
atcaactacc	ataagagcag	attgagttcc	tatcctaaag	gaaaatacta	ctattagata	840
aatttttagaa	atttatctaa	taaagctata	acaatcggta	aagctcataa	attttcttac	900
tgtgaactta	tatatgaatt	taaaatggaa	aggtattatg	tactaattct	tgtcccgatc	960
agaaggttta	cctgtacccc	tcaaataact	atcaccagaa	tgaaggcagc	agttacatga	1020
agaaycacia	agaaccacca	ctaaagatct	aatgggtaccc	aatccctttt	aaagtatgtc	1080
tgtgtctata	ttggtttaacc	ttttcttata	tgaaatacaa	atgccacatg	atctcactta	1140
taagtgggaag	ctaaacattg	agtacagatg	gacacacaca	aaaagaacaa	cagacacagg	1200
ggtttacttg	agagagaagg	gtggggagaag	ggtaggggtg	gaaaagatac	ctattgggta	1260
ctatgctcac	tacctgggtg	atgaaatcat	ttgtacacca	aactccagtg	acacacaatt	1320
tacttatgta	gcaaacctgc	acaaaacaac	cc			1352

<210> 98

<211> 913

<212> DNA

<213> Homo sapiens

<400> 98

ggcacgagtg	aatattttta	aggctcttga	tttgcctggag	gactgaaaaa	aatgaagtga	60
tagtgtctga	gaatatcat	ttgacttatt	ttttacagca	tccattccct	ttcatgttgg	120
gagtggtctc	tttagtggct	taaatctctt	gcctgccttt	gggagtgtgg	aggggtggagt	180
ggaccttttg	agggctcagg	gtgaatgtgg	ccttgcctgt	tggatagcct	tttgtttgga	240
ttctggctct	gggcacaggg	aataacacta	ctttctgagg	acagtatcag	gattgtctgt	300
agttcctgtg	agcctgaggt	gctgcatgtg	cccacccccg	tgtacaggcc	ctgccccagc	360

cacagcccac	tcaccttttg	accctcctgc	tctgcctata	cagtttgaat	accagcaggc	420
tcagctggag	gctgagatcg	aaaacctctc	atggaaaagt	gagcgtgcag	acagctatga	480
cagaggggac	ttggagaacc	agatgcatat	agcggagcag	cggaggagaa	ccctgctgaa	540
agatttccat	gacacctaag	ttgggatgtg	gatgtgccgg	ggtgaggaa	atgtggctgc	600
aaggctctcc	ggctgcata	ctgcatgctg	caggctctgc	ctttcatgac	cccaggcaac	660
agccagggcc	ccactcctga	gagacactgg	caacacctct	tagttgattt	ctgttttctt	720
ctcttttcac	tttttgttct	taccagggtg	gaggccatgt	tgaactggcc	tcttttcagg	780
acttttattt	ccccctggat	ggttgttggg	agggagggaa	agtgttttct	gaatggctat	840
taatagtatt	agatcattac	aacttatgta	actttcaaag	gttgtacaat	tatacaaaaa	900
aaaaaaaaaa	aaa					913

<210> 99

<211> 721

<212> DNA

<213> Homo sapiens

<400> 99

ggcagcagct	taaatacact	tttatttgct	atttctgtga	agcaaactgc	cgttggactt	60
cccattaaac	aatgtctaga	atctttgcca	ttgtctctct	atgctagtca	cacttacagc	120
acctactttc	agactcctca	tctttcaacc	tggtgtctcc	tttcagttgc	tggattaaat	180
atcattttgg	ctacatgttc	ccccctgtc	cctaaaattc	cttccccattc	caccccagct	240
attctaatagc	atttaataat	atgtctttgg	atatgtatga	tttcttcaat	gtacagtggg	300
atgtattcca	atttacataa	atagttctac	gttattttct	gttatggccc	ttcaatcaat	360
tccaagtttc	accttattga	tctcattctt	tctttccact	cagtgcctaa	gatgtgtgta	420
caactatgaa	tgcatcctat	tcatggcatt	taactgcagg	atgggtgttct	agtattcatc	480
cgaatttccc	ttatctgac	cactagtgat	ggtcattgtg	tcaattaagg	taatgttagc	540
taatggatca	aacacgatga	gtttactaat	aggagtatca	gtcatttctg	gctggatgca	600
atagctcata	catgtaatcc	cagtgttttg	ggaggctgag	cagagggatt	gcttgacccc	660
agattggaga	ccagtctggg	caacataatg	aaatcctatc	tctacaaaaa	aaaaaaaaaa	720
a						721

<210> 100

<211> 645

<212> DNA

<213> Homo sapiens

<400> 100

cccccccccc	cccccaagac	tgcaatgaca	aatgtctaca	caacaccgag	gtcgggggaga	60
cgcgagcag	aactccagaa	atgcctgccg	tgtctgcgtt	ctttagcctc	gctgcgctgg	120
ctgaagtggc	agccatggaa	aatgtgcaca	gaggtcagag	gtcaactccg	ctcacccatg	180
atggacagcc	aaaagaaatg	ccgcaggctc	ctgtacttat	ttcctgcgct	gaccagtga	240
gcgccctttc	attgtaaaac	attgtgcttt	acctactacc	ctagccttgt	ctttaccgag	300
ggatgctagt	gagtccaagt	ggtggaaaat	atagactgca	aacaagtgtc	tggtgcccc	360
cacggccccg	attcacttga	agcagaagt	agcatcctgg	gccagtttgt	tctctcagaa	420
cccagaatct	ttgagggtaa	ggttatctgt	ctgatactga	gcagaaacag	aatgatcctg	480
gagctttgct	ttctattgaa	ggcttttgac	ggtaaatagg	ggtaacttgg	taaaaggctg	540
cctttactgt	agctcaccca	gcattctctt	taccaaccag	agagtgtgaa	actagtttca	600
tatattacct	agttattctt	tcaaaacaaa	acaaaaaaaa	aaaaa		645

<210> 101

<211> 563

<212> DNA

<213> Homo sapiens

<400> 101
 ggacagagat aagatcgct taataccaga aatgattaga agtgctgatt tagattcaac 60
 aaataccata tgtccttata attttttgta agaagaaatt ggtaagtcc taactttcaa 120
 tgtgtacca aatacttgta tttatgctt tgataaaatg tttttcagc attaatcac 180
 atccgattat gccttattta tatatgaaga ataaagttac catgtttatac tgttatgtcc 240
 taaaattcaa atcactattt gagaaacct caaattggtg ctttcattat ataagtatac 300
 atttagacaa aaccccaaac taagccattt gaaacaagat tctctccatt gcagtttgta 360
 gcaatgttat ttctgtgtat gtcattgaga ggctaaatat cagtgttaat ttctgtttg 420
 aatccgtgaa atcatgcctg taaagcccaa acatttgtaa caaactccct aataaattta 480
 gagaaagtca aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 540
 aaaaaaaaaa aaaaaaaaaa aaa 563

<210> 102
 <211> 1324
 <212> DNA
 <213> Homo sapiens

<400> 102
 gacagactgt tttttgcaag accacattat attactttat tttttctgc tttttctttt 60
 aacgacatta gtgtttttga tcaactatatt ttaaaatgct ttttgtgagc cttttgggta 120
 tgtggaatct gtcccttagc tctgattttt tttctttatg gagcgtctta gggtactaca 180
 tgaaggtaag actgccacag tccccaggg aggcacactg tgttttactg attgatttga 240
 agatgataga gagcctaggg ggatgagtct attggactca aagggttacct ttgtttttc 300
 catttaattt aataatcaac aaaacgacaa agtcagttta atatctttct tctgctgagt 360
 cttcaggatg ttagctagtc tttcaaaagc cattgctaca aaagtgcaaa tttctcattt 420
 cctgtgggtc tctasaaact ctctcaatat ttaagccaa acaaatccat cttttctgag 480
 agacagaaac taaaaaattc agagttaatt cactattaaa aacagcaagg ccctgctatt 540
 tccaagaat aaagaaattt aaaattctca ttgtaagaag tatgaacttg aagctagaat 600
 tgggtttattt ctggcagcta cttacaatat taagaacttt acttttttaa tttgagacaa 660
 tttcacagga aagcaagaaa ttagtctatt gtgaatgtt tttacaacta acttgggaata 720
 tactgcaaat cactagtaag tagagacaca ttaagataa aatgattaat aaaacaattg 780
 gtaattctga ataaggatct gaataaatcc agatcacaga ctgttctttt accaatatgt 840
 aaaagaatgt gcaaaattmg taaaaatata cccaaaaatt taccagcaac agagaagcaa 900
 agaattggagt tcaaccttac actataaaca tctaattagat atgtatgaaa atttaaaaac 960
 tagaaaagggt tgggcatggt ggctcatgct tgtgatccca ggactttagg aggccaaagg 1020
 gggaggattg tttgagccca ggagttcaag accagcttga gcaatatagt gagaccctgt 1080
 ctctacaaaa acaaacaaac aaacaaacaa ataaataaat aaataaaatt agctgggaat 1140
 ggtgatgcac acctggagtc ctactactc agaaggctga ggcaggacga tcaattgacc 1200
 tcagaagtct gaggttacag tgagctatga tcataccagt gcaactccagt ctgggtgaca 1260
 gaacaagacc ctgtctccaa aaattaaata aataaaacaa aaaaaaaaaa aaaagggcgg 1320
 ccgc 1324

<210> 103
 <211> 1731
 <212> DNA
 <213> Homo sapiens

<400> 103
 cccgggtcga cccacgcgtc cgtaaaattg aattttgcag actaaattat tttataaggc 60
 tgttgctatt gaaagatcaa ttttgtcttt acctgcattt aaaatggaac agctcttagc 120
 agctgtggtt tttttttcca tttttttttt aaatttggtt gctcttaaga tgaataaagt 180
 ttataggtgc atctgcctcc tcttctctaa gaatatgcat actaatgttt gtttctataa 240
 atcaaatata catgtgataa tatgcatgta aatacataac tatcctcatc atgtatggct 300
 ctcaattctt aattgatgat aacatgattg aattgccata tatctgtatt acttgtgtga 360
 atggttgtgt tgaacacgct aagaacaatt tgaattttat taagtacaaa aatccttaac 420

gtttgtatgg	gaacattggc	aagaaagaat	agtgtgaaag	ggaactgttt	tagtttccta	480
tcagtaattg	tacatgcagt	taaatgttta	aggtaaaatg	attggtctct	gtcacagcta	540
aaagatttca	gtagccttca	ttgagtttgg	gtaaaaataag	ttgctgttct	tttgtcttct	600
ttttaatata	taaaagtatt	atttaaaaag	tatataacat	actatatatg	ttacatacta	660
acatatatat	acacaacaga	aagtttattg	gatttagtac	tgatattttac	tgtctatttt	720
tgttactttg	ttcttagtaa	tctattatag	ttacctattt	taaggtaagg	tctgtaggca	780
agatatgtaa	gtaaaggcaa	gccctacaat	tttaaaataa	caaagcttat	gtcttaagtt	840
gtattttttc	aaaggttcat	gtttttctga	gtaaatgtgg	tttatttagca	tgaaatatta	900
tgccttttac	ttaaattatt	ttatgtaaaa	tagggcactg	tttaattatg	aaagggggaa	960
aatcattcca	aataagagtt	taatttttat	taattataaa	aaactctgta	ttagtttcct	1020
agatgtgcta	gaacaaatta	ccacaaactg	ggtggcctta	aacaacagaa	atattattctc	1080
ttacacagtc	cagagactaa	atgtccagac	tcacaatgtc	ccagtgccat	gcttcctcct	1140
taggccttag	gaaagaatac	ttcctagcct	cttctcggct	tttgggtggt	gccagcaatc	1200
cctgtgttcc	cttagcctat	agtggcctga	ctccaatctc	agttttgttg	tcaagtggtc	1260
ttctaccctg	tcttctatgt	ttgtatccgt	gtccaaactc	ttttcttagg	gagaccagca	1320
ctggattaga	gttcaccatg	atccaatatg	acctcatctt	gactacatcc	gcaaagaccc	1380
tgtctccaaa	taaggtcaca	tttacagggt	accatgtgtt	aggatttgac	atatcttttg	1440
ggggacacaa	ttctaccac	tacaggccca	tatcttatct	aggatgttgg	tggaagaag	1500
gtgaagatta	ctacctcttt	tgatataact	agtttctgag	gtatttaaaa	atttggtttt	1560
aaaaatatta	agctttttgc	tcatttgcat	gtatactttt	ctctcaacat	tgttttgggt	1620
tatttaggtt	atttggttaa	acttcagtaa	atactaaagt	tacttgtatt	agaacataat	1680
aaaagtagag	aattaaacta	cttaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	a	1731

<210> 104

<211> 1466

<212> DNA

<213> Homo sapiens

<400> 104

ggcacgagct	ctacctgaat	gttcccctag	agtttcatac	acaatgtgtt	ggaaacctaa	60
atgtatcctt	ctcctcagtt	ttgtatttca	gtgtgtggca	tcatacaacat	ttgacccccct	120
aggtagtgag	agaccttgga	gtcaacctca	atgtcccatc	tccttccctc	tccttatcac	180
aggggtgtgt	tggttctcta	tgtcccggtt	ctcttaaaac	cacctcttct	cctccgcctc	240
tacagacacc	aacataaatc	aagtttccat	cttcgtttgc	ctggacaagt	ggcaaggcag	300
cactgaaagg	atactccttc	ctctagtctt	ctctgccttt	tgccactactga	gcccactctt	360
ctgagctgct	gataaaggaa	ttcacatacc	acacatcctt	tgatgggatt	gccatgctac	420
aaagcagaac	ctaaatccca	tgccctggag	ttaggcagtc	tacattctgg	cttctgtgac	480
ttttggccta	atttttgcat	cagccccaaa	tttctgttgt	gccaccatcc	cagtggtatc	540
tagaatttag	tcttacacaa	tcattccata	ttcctttaat	gagtccttta	gcatttgctc	600
attcctttca	tggtccctat	ccccgtacct	ggaattactt	ttcctctttt	acttactcaa	660
gtcctgcaaa	agccagttcc	attatgctgg	tctcactgac	ctctttctac	atattttctgg	720
taagaatgaa	ttactttctc	ctgaaatacc	tctgccatat	tgtttaaaaa	ttgccatatg	780
gtgctggaca	tgagtatgtg	ttcacatgtt	tattatctac	tctagtctca	atttctaagg	840
tcttgaatat	aggaaccaat	ttattcatca	ccttattcca	gacatgatgg	aactcagctt	900
tattgagaat	caagtgatta	tagtagatag	tgaccatcct	gagtatgttc	atgtgttaca	960
taacaatgtt	ttggtcaacc	aaggactgca	tataggaagg	tggtgtcata	agattaatat	1020
ggagctgaaa	aattcctaata	gcttagccat	atcgtagcca	tgatattgta	gcacaatgct	1080
ttactcacgc	ggtgatgcta	gtgtaaatgc	tgccctacca	gtcatataaa	tgatagcac	1140
aaggggccag	gtgggggtgg	ttacacctgt	aatctcagca	ctttgggaag	ctgagggggg	1200
aagattgctt	gagcacagga	atacaagtct	agcctgggta	atgtagggag	ggcacgtttc	1260
tacaaaaact	aaataaaatt	agcctggcat	gggtggcagc	acctgtagtc	ccagctactc	1320
tggaggctga	gacggaagga	ttgtttgagt	ccctggaggt	tgagctgcag	tgagccatga	1380
tcatgccact	gcactccagt	ctggggcaca	gagcaatccc	ctttttcaaa	aaaaaaaaaa	1440
aaaaaaaaaa	aaaaaaaaaa	aaaaaa				1466

<210> 105
 <211> 1303
 <212> DNA
 <213> Homo sapiens

<400> 105
 aggttaaatg cgtacttttc taacctttgt tattttgaaa gttattctga tattcctatc 60
 cagttgtgcc tcatctacta gaaatttgct cacatggcca aatgatgtat ccacagaaca 120
 atttgaaact agaccttttg gaagcgaact cctacaaact gtcatacatg ttagcagaac 180
 ttgagcaaaag acctcaaccc agccatcctt gtagtaattc catcttcagg tggagggaaa 240
 aggtaacatt taaggagact gggtgtaatt tcttgattgg gcctgctggg tggagtggct 300
 taaagtagca tcagggcaaa aaagggtgta ggaattctat gtgataataa tattcatgca 360
 gttagttaaag aagataaatg ttttwatttt tcttttgagc acaataacaa gagctagaca 420
 aaaccgaata cattctgtgt acaccaaact tctatgagaa gctaaaaaac acttttgatt 480
 tcttctttct catcatacct gaatttcac ctttggatgt gcttttacag taaaatttct 540
 attaaattga aattttaata ttcgttcaga cctaaattat aagattttgt ggtatgtatt 600
 agtctcatct gttaaagatg gtgcctaata cagataatgc atcagtacag ctctgaaatg 660
 cttgtagcta tttttattac tgatcagaag ggggaactgt aatcatcttg tgaagggaca 720
 gttttctaag gctcaagagc tcgaaaacaa tctcaatcat ttacaggggt gtgatcattt 780
 cacttgcat aagccaacta aagttgtatt tgtaaaagta atgctatgaa tattactatt 840
 tgacctagac acataggtta gaattggaaa cacaggctat aaagtatagt aattgtgtaa 900
 ttgtgaaaat attaaaggct caactcaaaa ctgaaacaca gtagggttta gaaatctttg 960
 aattatttat acccctcagt ttaaaaactt ccagtcagg cgagtggtct catgcctgta 1020
 atcccagaac tttgggaggc caaggcaggc ggatcacctg aggtcaggag ttcgagagca 1080
 gcctggctga cacggtgaaa ccccgctctt actaagaata caaaaattag ccaggcatgg 1140
 tgggtggcac ctgtaatccc agctacgggg gaggtcgagg caggagaatc acttgaaccc 1200
 gggaggtgga gggtgtagtg ggccaagatc atgccactgc actccagcct ggggtgaacag 1260
 ggcaagactc tgtctaaaaa aaaaaaaaaa aaaaactcgt agg 1303

<210> 106
 <211> 1516
 <212> DNA
 <213> Homo sapiens

<400> 106
 ggcacgagag gaattgatgc ttttaattttt ggatactttt tcagaatttt taatttacta 60
 tgggtccggcc taagatcctc tgttgcatca ggttttgtgc acaaaagaaa agcacaaaag 120
 ttgaatgcac atggggcatg tgctttctgt gcaccaaata tctggatgag gttctttttt 180
 caggcctaca gtcaaatctg tgtccagaat tttttgactt ttttgctttg tataatcata 240
 gaattcattg ctgctgattt ctataatgat tcatgttgct atgtgtctct taataactga 300
 gggctgtcag taacctgtga ttttgccctt tctatagtct tactcccatg aagaaccttg 360
 gttctgatgg agaaagtga aagctttatt tcttccccta gatattctta ttttctatt 420
 atatttttta gttgtgtact gtgtactaga gatttttttc agtttgttat gaacacaatt 480
 tggtaagccc taaattgggt ctgcctgtct ccaaacagaa acatctgtac aaatcttggt 540
 ggtatagact actttctgga aaatgggtcaa gataagttca tgttttcttg aaatttctaa 600
 gatagtatat ggtatcactt gtttaaagca aatcagactg agtttgacat ttaattcaat 660
 atttctggta ttcagtaacg ggtatatatg tttgttcttc cagtttgggt cagtttaaaa 720
 gatattgtgc aaagtataca tagaaaatgt gagcaatgcc tctctttgcc ttttgatcag 780
 aaacttcagc agagcggtaa ggattccaca tgatttaaac tgaaatgctt ttctttgttg 840
 ctgtaagaac ttaaaatgta aaataccttt ttcagtttaa gtccctgtaa caacattgaa 900
 gcatggagat gaggcaagga atagtactca ctgaagtga aatgactgcc cacttcaaaa 960
 tcttcattgt gtttacacac cagtgtattt atacaaatca gaggcatttt gtagatgctt 1020
 tgctgacttg ttcagctctg taaaaacaca gaaatcagac ccattttgta aagcggaaaa 1080
 tcatgttaca tgggaacatg cctgtatata tcacatacat ggtaatggag tcttaatgat 1140
 aagtgaaga taataattta atgatgggat tagtctgac gcttaatatg cacaatcctg 1200
 gaagtgaatt acttgcacat gatatagtga tatttattat tctgtacaga gagaaaaata 1260

catataaaac	atatgcttac	attacatgca	cgcggtatttc	atgctccata	atctttttcta	1320
ttttttaatt	tacctttctg	taaatgatgt	gcatggaata	tgctttatag	aaaaatgctg	1380
ttcataat	gactacgtgg	aaaagtgcct	atatggtggt	aatgctagta	aggcaataa	1440
gacaaattat	catgttggtt	tactacatca	ccagtttaaca	ttttatattg	tgatgtttaa	1500
aaaaaaaaa	aaaaaa					1516

<210> 107

<211> 1689

<212> DNA

<213> Homo sapiens

<400> 107

actatagaag	tcgcctgcag	taccggctcc	ggaattaagg	gtcgacccac	gcgtccgggc	60
taattgtttg	gtcagaaatt	cctaaggcca	cagctttggg	gggttcgtgt	agatgtacat	120
gggtgggtggg	ttataaatat	tgggacttaa	ggcagcttgt	tctatgtatt	tatctttgct	180
cttgggtgac	ttagggaatg	attttatttg	atttaacctt	ctttctgttt	gccccgagaa	240
tactcgccag	tggcgcttgc	agttgtagca	tttaccctaa	gataactttg	cctacgaaat	300
atctcgcttt	tattatttyc	acatcattct	agtatatgga	ctttggaaac	aaaagacatt	360
gttctattta	tagcattctt	tttttttttt	tagtagcggg	atttccattt	acaaaatata	420
gtaactcttg	attactgaaa	atgtcaaatc	ctagaaaacg	tagcatgcct	atacatgatg	480
ttaacatcat	tctcgaaacag	ttgttgccg	aagattcatt	tgatgaatcc	aatttttttg	540
aaatagacaa	ttctgatgtt	ctcttttagaa	ataactcagt	ttttatcttt	tttcacattg	600
aaaatcagtt	agatttgctt	aagcctcaaa	gagaatgttt	atgtaaaatta	gcgctggcaa	660
tttttttttt	tctaaacagg	aaaagggtta	aatgaagggt	gataaaatgg	atgttcaatt	720
gtctttctga	aagtgtgagg	cttgaaggga	tgaataaata	ttttcttaat	atattcaaaa	780
aagtgcattg	ctttctgtga	tggaaagtta	gacctaaatg	tctggaagtt	gtaaccctca	840
acacagcttt	tcctgatttg	ctgcaaaggc	acatagctga	ttatagaagt	gaagacggca	900
aggacgggga	ctccaacaaa	ggaaaccttg	ttgcaggatt	tgggaacttt	catgcttcag	960
atgaaattca	ggcatgtgag	catcactgca	gaatgtgggt	catcattgcc	atcatgagta	1020
atcacttgct	gtcctacttt	ctgagaccaa	gactcttttg	tcataattct	tagcaatagg	1080
acgggtaaaag	actggattta	attgctgttc	agagtataaa	aactcaattg	attccaacat	1140
atctgaatgt	gcagtaaaagt	cttaaaagtc	aaccgttaat	cattaagtct	tttgccctcta	1200
aagtcttttg	cctctgaaga	agtttattac	atgagttgat	tttcataatt	tcattttggt	1260
gggtttttcc	tggtgttggg	caagggtggg	tcacaggaca	tgggactagt	aagcatttta	1320
ctgtttacta	tatttgtctt	tttataaaca	gtatctccca	aaatgtgatt	agaaggctac	1380
caagcctgta	tttggacatt	taattgtgtg	ctttatataa	tgtaaactact	aacagratatt	1440
ggactgcctg	ttcattcctg	gagacaaaaa	tgaaaatctg	tcagttcaag	ttcttgggta	1500
acatcaagtc	attagaattt	atctaaagct	tatcatgatt	tgataagaca	tccattgcat	1560
gcagctgttt	tagctcagtg	caaaacactg	aaattgtgat	tccttagactg	tttctgagac	1620
atctggatgg	aaataaatgt	ataaatgtta	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	1680
agggcgggcc						1689

<210> 108

<211> 1943

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (161)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (1926)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (1928)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (1934)

<223> n equals a,t,g, or c

<400> 108

ggcagcaggc	ttggctaagg	tccgcgggaa	cccgtgagcc	accgagagag	cagagaactc	60
ggcgccgcc	aacagcccag	ctcgcgcttc	aggtcccggc	gccgtcgcgc	actcctccga	120
tggccacaga	tgtctttaat	tccaaaaacc	tggccgttca	ngcacaaaag	aagatcttgg	180
gtaaaatggt	gtccaaaatcc	atcgccacca	ccttaataga	cgacacaagt	agtgagggtgc	240
tgatgagct	ctacagagtg	accaggagg	acacccaaaa	caagaaggag	gcagagaaga	300
tcatcaagaa	cctcatcaag	acagtcatca	agctggccat	tctttatagg	aataatcagt	360
ttaatcaaga	tgagctagca	ttgatggaga	aatttaagaa	gaaagtccat	cagcttgcta	420
tgaccgtggt	cagtttccat	caggtggatt	ataccttga	ccggaatgtg	ttatccaggc	480
tgttaaatga	atgcagagag	atgctgcacc	aaatcattca	gcgccacctc	actgccaaagt	540
cacatggacg	ggttaataat	gtgtttgatc	atttttcaga	ttgtgaattt	ttggctgcct	600
tgtataatcc	ttttgggaat	tttaaaccct	acttacaaaa	actatgtgat	ggtatcaaca	660
aaatgttgga	tgaagagaac	atatgagcac	atgagttaag	attgtgactg	atcatgattt	720
atttgaagat	ggagcactgc	tgattttatga	aggaaaaaag	aagaattttc	taaagattac	780
acataattca	gaaagacttt	acccaattca	gttgtcagac	ataatgattt	atttgaaggc	840
ttgtttttatt	tgaagaaaag	catattgcc	aaaattcttg	ttaaaagctt	cctaaccgggt	900
aacagaccat	gggagagata	tgtggttggg	taatgcaaat	gtagttatac	aaagaaaaat	960
acagatgtct	ccagacctga	ggacttttta	atagggcagt	tggtgtgttg	gtggcacatt	1020
ggatattttct	aacatgtaca	aagctatgta	ttttgattta	ctttcatttc	ttgctatgta	1080
tatgtacttt	tcttaaaatg	ccaagaactt	tctcttgcta	tcattgctcc	ttttgaaaca	1140
attcaatttt	catgtctaca	gctgactgtt	ttgttaagat	tgagtcatcg	acattcagga	1200
tttaagtctg	aggtagtcaa	ccctcaggaa	aaaaaaaatg	gcttatctga	aatcagtact	1260
gtggaatga	actatattag	ctattatgaa	taatgtccag	tataagaata	tgcttctgga	1320
attgagttct	ccttttaagt	accaatgata	cttaaatctc	tcagaaatgt	aatgggtgtg	1380
cattgccttg	aatgcttgc	ttagggcttc	ttttatgtta	tcttaaaaag	tgctgggtgaa	1440
ttttccattt	tttacatcca	tttcacatgt	aagagacaaa	aaagtctaga	ttggctcttga	1500
tattgagata	ataaaaagta	agtagcatta	agaaaagtaa	caatcttcat	tctacagatg	1560
aactcattga	aacaatttag	gggaatgagg	ggcaaaaagg	gagaaatact	gctaagaac	1620
atgagcataa	aaatgcgtgc	gtttcagtgt	ttaagaaggc	ttgataaaga	atgtcacttt	1680
tttattttaac	tgataagatt	tttgttattt	tttactttga	taagtaaacc	aaagaatatt	1740
tgtattttcaa	gcagttttgtg	tggtgtttct	atataatttt	ctgtgtataa	ataataaagt	1800
aggcattttgt	ttattttgta	aaaaagaaat	gaaaaatctgc	tgggccagcta	tgctcctctag	1860
gaaatgacag	acccaaccac	cagcaataaa	catttccatt	gtcactgtaa	aaaaaaaaaa	1920
aaaacncngg	ggtnccttttg	ggg				1943

<210> 109

<211> 1594

<212> DNA

<213> Homo sapiens

<400> 109

ggtggattat	ctataaaattt	tggaataaat	gactttttcta	ttttttttct	tattttccag	60
aatactttgc	attaaaaatt	tagatcttct	tacatggaag	aggtcaaatc	cagttatagc	120
aaaacacctt	tactgcagag	gccatataac	caaaaagtcc	aaaggcccag	cccagtgagc	180

catttacttc	agtgatgttc	agtacaaaat	ttcactgcc	ttaaagactt	tggaaagtc	240
cttttaagtt	gtaacaggat	ttgactatta	gtgtggattt	cgataaccct	tcttttagtta	300
atgtaataaa	tgctctttga	agctggaaat	atactgggtt	cttgcaccct	tgtagttttc	360
tctgggcttt	atgcaaaaga	tgtattttct	gaatgtctta	gtatgttgct	gtctttttac	420
ctgcctgggg	ttgccttcag	atgtgggtga	ggacttggtt	acaaaaagtc	ctagaagacc	480
ttaagttaaa	attagctctc	caattttcaa	gtgttttctt	ttggcagaat	aagttttggg	540
tcctatgaga	ggtaaatagg	ttggktgcct	ctcttgctat	agcccatcta	gagcagaaca	600
aaaactttgt	ttttgagggg	gatggagtta	ctttttacct	tgtcccctaa	ttgaatatat	660
ttaggacttt	tggttttaga	aatctcagag	ttctatagtt	ggaatagtgt	cacagagata	720
gtgttaggtt	tcttaggcaa	caaaattgga	gtgttaccta	ttgcccatct	gggcatggct	780
ttgatgagtt	ctatataaat	aatgagtgt	gcctaatttt	ctgtcatgtg	ctacctagca	840
tatacatttc	tgacagtcag	gtggattatc	ttgaaatttt	agacaaactt	tctgcctcct	900
tggaaagctc	tacatctcct	tgtgattgga	gtgtgggttt	caagattcca	gaargtgggt	960
gtmttcattt	tctattttca	atgcactcta	aagatgacct	gcttacaggt	tctctgaggc	1020
agacctcagt	tgctcttggt	atatgttaact	ttgctttgat	aaggaataaa	caattacata	1080
tccagataat	gatacagaat	tgtcaaaatg	tttaatcatt	aagccactac	aaatagacat	1140
tttcgtggaa	taaaactgatt	ttgaaatgtc	tttcaattta	tattgaaata	tctgaaggca	1200
cctctaaaat	ttgggtgatg	acacagaggt	atagattata	agatttaatt	atgaatgggt	1260
ggtataactgt	gaagactacg	gaagtgtgca	gaaagatatg	tttggctttt	ttggttttta	1320
gtagagatt	ttaaaaaggt	ctgtgaaggt	caagttgatc	attgggtttt	attgcatccc	1380
acaggtttct	ataaaactggc	ttaaaaatgg	catactccat	tttctgggtg	attcattgta	1440
acttttaggaa	gtattttgaaa	tggtcacatt	aaagtctagg	aatgtcaaat	tgtatttttaa	1500
tgcgtaggtt	ctttttcttt	ttcttttttt	ttggagacag	cctgggcaac	aagagcaaaa	1560
ctctgtctca	aaaaaaaaaa	aaaagggcgg	ccgc			1594

<210> 110

<211> 1742

<212> DNA

<213> Homo sapiens

<400> 110

ggcacgagct	cgtagcgctt	tgtagtctag	ggagttaa	taaagtaagt	ggagacaaaa	60
gtactctttt	gagagctgtc	atttctctta	gtgtgacgct	attaataatg	tagtgtaatg	120
ctatttttga	agtttggttc	tttcttttct	ttttgtcttc	ctctgactct	ttctgtatt	180
ctaaatgaaa	ggggaataat	gcacttagag	gggggcactc	tcctaaatc	actgtctcat	240
gtacgacatt	atctccgact	tcggctctca	tgttttgaaa	aaatacctct	tcategctct	300
atttttattt	ttcttcttct	tttattgtga	atctctttta	ccaaaaacat	ttgtagggtt	360
cttcacaaaag	attttttttt	tcaatcagga	tgaaaactag	atcatgatgt	gaccatttca	420
ctgtgagtgt	aacttccctt	tttgacagct	ccattagatc	tgccagggtta	taaattctca	480
tatttctgac	ttgccttgaa	atcagaaagt	gttttcatta	tgctagtctc	tgtgagcaac	540
aagcatgaag	gaaggcatgg	caggatcat	agccctttg	atgaacttac	ctgtttcaac	600
tcagtgccag	ggcagaacat	ttactgctaa	ccctgatggg	tcaactttga	ttgcaaatga	660
tgtgtggtac	attttgaatt	taaagaatgt	ttctgagatt	attctacgat	cacttgtcat	720
ttttatgtgt	gcagtaaatgt	gttgtgtata	acttggattt	caacaatc	cattgtttga	780
aagttagaaa	atattctaa	aatactaatt	atcttgcctca	aataatcatt	taagtacaac	840
tgtcacttga	ttatggtgaa	tatttttaag	taaaattata	tatttaaggt	gtgctacctc	900
taattttatt	gtcatacaaa	aagcagatta	ttgaacatgt	taatgtaaat	tgtactttta	960
attttttcca	gtactctaga	acatgtgtaa	ggttaaaaga	atttaaatga	cccagggttt	1020
tcttttttaca	taataaataa	gaagaaatca	caaaggaagc	agatattata	ttgtttttta	1080
tatacacatg	aaattgtttg	actttatttt	gagacctcac	acaagtataa	acatggcagt	1140
ggtgtgtatg	atcaaagtaa	gaaattaaag	agttaccggg	tctttataaa	ccagaagtc	1200
attgactttt	aataatgctg	tctcaaatat	ttgatagtaa	attgtggaaa	taatacaagc	1260
tgagcctatg	ggactgtact	ttgtagtact	gtttaattta	ataactctaa	taatccctta	1320
agaatattag	gaaaaatagg	ccgggtgcag	tactcacgcc	tgtaatccca	gcactttggg	1380
aggccgagga	ggcgggatca	cctgaggtca	ggagttcaag	accatcctgg	ccaacatggg	1440
gaaaacccat	ccctacaaaa	acacaataat	taggcaggca	tgtggtgag	tgcctataat	1500

cccagctatt caggaggctg aggcgggaga atctcttgaa cccaggaggc ggagggtgca 1560
 gtgagccaag attgcgccat tacactccag cctgggagac agagcgagac tccctctcaa 1620
 aaagaaaaag aaaaaagaaa aaagaatatt agggaaaaata tcttaatgca aaatatatta 1680
 attagtaatc tgccaacact gagatgtact ataaggccaa gaagaaaaaa aaaaaaaaaa 1740
 aa 1742

<210> 111
 <211> 1501
 <212> DNA
 <213> Homo sapiens

<400> 111
 ggcacgagcc tatttctgct tactgtgtta ccagagagcc tgggggtctg gatcctatct 60
 ggccccgtca ggggtggattg ccaaatgagc agttctcttg cccagtgccc ttctctgtgc 120
 tataaataag ccccatgttt attttcttat gttattgaaa tgagcacttg tgatttgggc 180
 ctctttttgag gagtccagag agcgtccatc cgggtgcctgg tgagggccct gcatggctgg 240
 ctgctgtctg aagctatttg gagtcccttc cctgtgtttt ctatgtggct taatttcaat 300
 agaaaggggtt atatgcaacc ctgtatctgc tgattttcag gtttcaactt tctgccagcg 360
 tcaactgctg cttagaagta aagttatgtt tcccataagg ggataacagc cacaattgag 420
 gtaattaacg aaaattgtac attgggtggca gcacctccta taggatttcc aatagtcttt 480
 ctctagtaga tcattggggg ctcaccttga tctctctctt tctgtctacc ctgcacaaa 540
 ataccttgtc ctgttttctg gatatagttc caataatttt tttcctaaca gcctttttgt 600
 caccagttgg ttgatattct tacaacttgg ccaaatgagg gtccatttaa ctccatcttg 660
 tctaattgcat ggagaattca aggatttttt ttttctctt tccatagcac cttccagttg 720
 ccagttgtac cctggccctt ctttggaggt cataatgatg aatatccatt aataagagat 780
 tgatgctctt tcaactctca tgtcatctat accatctcag tggagaggat gactttggat 840
 gaggttggaa tacaaggaa acatttggaa gtccactgca gtgtattata tgctgtgtgg 900
 aagtctgggg gttaggaaat acctggaggg agaacttctt aagaaatgat ttttggttct 960
 tttaggcctt aacagcaca taaaagtatc ccatgagacc attatgagca ggacaccaga 1020
 ttgtttcaca ccttgggctg tgactattta ctctcggta cagattactc tggttaaatc 1080
 actcagtaaa gaaatctttt catgctcaca atctgaacct gaaggctatt actgaagaga 1140
 attgcatctg acaacaaaat ttaatttact tccagagaaa ggaccagaag aaagttaaatt 1200
 ttcatttatg tttttaagtc tattgtctta aaaagattct tttcccttaa aaataaaaaa 1260
 aacctgatgt gatgggttcc ttcagtcaac aaatacttat tgagcagtta ttgtgtgcca 1320
 gatactgttc ttgtgtgag gatattggcac tgaacaaaac aatgtacctt ctttcgtcaa 1380
 gcttacattc tagtgaggaa gataaccaa acaagtgact gaataataatt tcaaatgtca 1440
 ataaatgctg tgaagaaaat aaagtcagag tattatatgt aaaaaaaaaa aaaaaaaaaa 1500
 a 1501

<210> 112
 <211> 791
 <212> DNA
 <213> Homo sapiens

<400> 112
 ggcacgagct gcatttgatc tcattcttta gtccaatgta agtaagagta aaacaatgac 60
 atttaaggcc accaggctat tctcattttt ggaaaaatgc tggattacat taccagcata 120
 ttaaattgaga atatcaagggt gtaatatctc cctagaaatt gtctcacctt caatactatt 180
 gacatttttg gacctgataa ttttgttggt ggctctagcc tcatgttata ggaggtttac 240
 cagttttcct gccctaaact taccggatgt gaatagcaca ctccactacc tacagcagta 300
 aaaactaaaa ttgtctctaa acattgacaa attgtccctg gtatgaaaaa tcaccctgg 360
 ttgagaccgt gttgttgaaa ataaaacaaa aactttcaca tcaataaata tgttaggctg 420
 tgtatgttaa ggattaacat taagacaata tggagcaagc actacatgaa agcagtgcag 480
 attggggaat tagtggcaca ttatcctaag agttaatata gtgactgtaa tatctaaata 540
 tcatcctata gagtttttct tagatttttt cattagtata acaggatggt gtgtatgtta 600

caactgtatat	actgttattt	tgagagacaa	ttttgggaat	tttgccaagg	tattttcaat	660
tataggtott	taatacattc	taagcaagt	ggtctcaaaa	atgggaattt	tacacccac	720
attcttcttc	ccatccggtg	gacatttgtc	aatgtgcgca	aatatttctg	attaaaaaa	780
aaaaaaaaa	a					791

<210> 113
 <211> 1637
 <212> DNA
 <213> Homo sapiens

<400> 113						
ggcagcagca	ccactccctg	ctcttctgca	ccccaaatct	tctttgttgg	gaaaagaggt	60
aggagggagc	tggctgggag	gctcctagtc	tggtgggaag	cagtggatgg	tggctcctct	120
ccatctcttc	atccctttct	cttggctagt	gaggacaata	gggcaattac	tgagtcctgt	180
gggcaaggca	ctgagtcac	ggtcgaatca	gatgatgccc	aggctcctgg	gatgagtgag	240
tcactcttaa	tgggcagctc	ccaagatgaa	tggtgagagc	atcctgcctg	gctttatgcc	300
tgcaagccct	ccccgtaac	tccttccttc	ttgcagggtg	gcaggaagaa	gcagggtaga	360
aaagtttagat	tcctaatact	actcctaccc	ctcaacccca	agggaccttg	ttggtaata	420
gcgaagggaac	tgggaaggat	gttcaaaggc	tgaggcaggg	cacagatgtc	acatttcac	480
tctgtggaag	gtgggctgct	caggccagat	ggatgagctt	gtttgtgtgt	gaatgttct	540
ctcaccttcc	tgatggtgag	gggggagatg	tccacttcca	gatgctgcca	agtagacttc	600
ctgttttctt	cctcttgctc	ccccccgct	tcctttttat	atttacaag	ctctctggtt	660
atgtacagca	gggaaatggt	gcctgagaga	ctccttcaga	tagcagttcc	ttctagtgtg	720
agtcaggagg	caactgcgtcc	ccagagtccc	tgcatcctca	ttcatgaaat	gctggcagtc	780
aagggaacag	ccctggctat	ctcatgaggt	tggccttagg	atttaagtga	gataatgtgt	840
ttgacaatgg	aatcccagca	tgagcttggt	tttcggcttc	attaaaaat	tagaatttat	900
tgtagaaaaa	ttagaaaaata	gacagaaaaa	ataatcctag	taccttctct	tactatttga	960
tgtatatccc	atattttttt	caaaaaaaag	aaataatgat	aataacatga	tagtcatgta	1020
tcaatgcatt	ttctactttt	gatggttaca	ttattattat	gtgagagaag	gtcttatttg	1080
tggagataac	aactacaaaa	agtatttggg	gggggtgatg	agcatcaggt	actcaactca	1140
atggttgaga	gcaaaaagtt	atttgtactg	tactttttct	taactttgtg	aatattttga	1200
aataaaaagta	cgtatgccat	ttggagaaaa	gaagatagca	tagtggtggt	tttttatgtc	1260
tggaaatatcc	atgtctgtct	attccagtgt	ctggaatata	catgtctatc	tatccactgg	1320
aatatccagt	gtctcagttg	agggtaacta	aagggaaaaa	cctcccacac	ttgactttct	1380
ttgaagagtt	tacaaagatt	tgtaatttca	gcctgggcaa	catggtgaga	tcttgtctct	1440
atatgaaata	aaaaaaattt	aaaattagct	ggatgtagt	atatgtgcct	gtgggtcctag	1500
cccttaggag	gctgaagcca	ggaggattgc	ttgagcacag	aagttcaagg	ctgcagtgag	1560
ccatgaagtc	tctcttgat	tccagcctgg	gtgacagagc	aagaccctga	ctcaaaaaaa	1620
aaaaaaaaa	aaaaaaa					1637

<210> 114
 <211> 1588
 <212> DNA
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (778)
 <223> n equals a,t,g, or c

<220>
 <221> SITE
 <222> (1150)
 <223> n equals a,t,g, or c

<400> 114

gggaacggtt	tcccagctga	accacatgtg	acgaaccaac	ccatttatag	tctaaaaatt	60
atgagaacaa	gtctcttttt	tttctttttt	aaaaacatct	tggtagcttg	tggtagcttg	120
ctcatctcca	gaagttccca	ttcccagagt	gccccgcgag	gctgttggtg	gccacataag	180
tgaccacagc	agggcctgga	atgcaagctc	ttttcccggt	ggctttggag	cagagggggag	240
tcagctgggt	ggccttggtc	cctgaggggc	attggccagg	gtcacactcg	gttttctagg	300
gaccaggctg	ttccatgggg	tggagaggca	gtaggccatc	cctttctctt	catcttttat	360
ctagcccttc	cccatacaaa	cactttctca	cagaaagact	tgtttggtgt	cccctgagca	420
aggccacagc	tctccacact	atcttctgtt	caggttggca	gatgggttca	ggcaagggtc	480
attctgagag	gagacctcca	cccaatgctc	ccgtctgctg	gcttggtcac	cctgagctgc	540
tgccctgagg	gccaggaaag	tggctgggtg	ccccggcctg	tgagagatgg	ccagagcctc	600
agccagaccc	accctgccag	gagaaaaggg	gagtggccgg	agaagccacc	ccttgagagc	660
aggctgggct	cgcccttggt	cccaggaggg	ctgcacaggt	gcagttgcca	ctttcagacc	720
acacacttca	agctagagct	tggcctgtgc	ccctgggtgt	tgggggttgg	aagarcanac	780
tgtgttgaca	gggtctktaa	ttcmcgattc	caggtgcttc	tgawttgttg	gacctgatg	840
aatcgcaatc	aggactatgt	ttggcttgag	ctccagcctc	atgaccacca	accccatcga	900
atgactgaga	gctgccaggc	tcctgccacg	ttggagctca	saatcctccc	gtggggacac	960
tgtgctgcag	ccagcctcca	gggagccgct	tgttttatgg	cctaaaaagt	ctctttatgc	1020
tgagaaggtc	acactggact	ttgtctcacac	agagatttga	cccaagaaaa	ggaataaaga	1080
kgccttggtg	accctgtctt	gcccctgttc	ctgccccgcg	cgtaccctgc	tggctcatgs	1140
cagaagamtn	ggaacttgct	gtctgtgggc	atgtttgttg	tgggctgamg	gtagcaccca	1200
cagcactggc	cgtgtgtgtg	ccacagtcag	gaacgargtc	ctgcccccca	cccattaacc	1260
cattagcttt	gtccttactc	cttgaggccc	tgcttgaaaa	gattaatggt	gctttgggar	1320
gctgaggcag	gaggattgct	tgaggccggg	agctcgagac	cagcctgggc	gacatagcaa	1380
gaccctgtac	ctccaaaaaa	taaaatataa	acattagtca	ggcatagagg	cacacacaca	1440
cataggccta	gctactcagc	tgaggtagga	ggattgcttg	agcccaggag	ttcaaggtta	1500
cagttagcta	tagtcatgcc	atcatacact	cagcctgggt	gacagagtga	gtccttgtct	1560
ctggaaaaaa	aaaaaaaaaa	aactcgag				1588

<210> 115

<211> 1926

<212> DNA

<213> Homo sapiens

<400> 115

taaccatggt	cttaacatct	agatgcagaa	atggatttga	atgtgatctg	cagtttctgt	60
cgtagttatg	aaatggcttt	tgccaacatg	ggttacccca	aagcaaacat	tgcattgctat	120
ggaaataactt	cttgtctctt	tttctcccat	ggtatctgta	ttgcttcttt	aataataaca	180
tcatgtttta	cagcctgcac	tctgtgccca	tgtttattca	acccttttta	ctttgggtgac	240
tgtttatcag	atgtgttgcc	ttgttcttcc	tgccctccag	agcccatggt	tttaattata	300
actgaaatat	gaaatgataa	tatttgggct	ttatttcttt	tctttaaagg	acagtactgc	360
tttaaagaga	cagtgttaag	gatcttggaa	gcacagccaa	catgtgtgtg	acatggaaga	420
agccactaac	caactcctag	atgtgaacct	acatgagaac	cagaagtctg	ttcaagtgc	480
agaaagcgac	ctcggaagtg	aatctgagct	tctagtcact	attggagcca	ctgtacctac	540
tggctttgag	caaacagctg	cagatgaagt	cagagagaaa	cttgggtcat	catgcaaaat	600
cagcagagac	cgtggcaaga	tatattttgt	catttcagtg	gaaagtctgg	cacaggttgg	660
aatggagcaa	tacttaaaat	agttttctaa	gttgatcatt	ttatggttac	ttttcatttt	720
atagaccctt	ctggaatcct	ttctgatttt	cattgtctctg	ccccaaatca	ggtgatttta	780
aagttactgc	attttctatt	gagtactgat	ttgtaatttt	gactgtagtt	tcagtatccc	840
ttaggaccaaa	aagtgtttca	gatttttgga	gggggggtggg	gggagtcttg	gaatatttgc	900
attatacttg	ccagttgagc	attttctttt	tttttctttt	tctttgagac	agagtcttgc	960
tctgtcatca	ggctggagtg	caatggcgcg	atcttggctc	actgcaacct	ccgcctccca	1020
ggttcaagca	atttctctgc	ctcagcctcc	ccagtagctg	ggactacagg	cgcccgccac	1080
cgcactctggc	taatttttgt	atattttagta	gaaacggggg	ttcaccatcc	tggccaggct	1140
ggtctcgaac	tcctaaccct	gtgatccacc	cgctctgacc	tcccaaagtg	ctgggattac	1200
aggcctgagc	cactgcgctc	ggcctgccag	ttgagcattt	ctaacttgaa	aatccgaaat	1260

gctgcactga	gcatttcctt	tcagcattgt	gttggggctc	aaaaagtttc	agatttttag	1320
atgtgggctg	ctcaacctgt	agtaggattg	ctacattttt	agaggtgata	tagtattttt	1380
tgaggattga	attcaccaga	aagaatactg	ttcaacttaa	gaaattctct	aagtattctca	1440
gtattttatt	acttcagcat	atagcagatg	tttaaccttt	ggtgatgaaa	cataaaagag	1500
attttccttt	taatttgcac	caaaaattct	tgctcagaag	taattaggac	tcactcaagg	1560
taaaaagtga	aggtcagaaa	actgcaatct	cactatcata	aagctgatata	atctcattac	1620
cacataagca	gtgagagggc	aaggattgtt	gtctattttt	ttttccacca	atatatccca	1680
cgctcttagt	gtctggcata	tagtaggtac	ccagaaaata	tctgtgggat	gaatttgtaa	1740
gcaaatgttt	tccatccaat	attttagcca	ataccctcag	cagctttgtg	aagataggca	1800
ctattattat	cctcattaat	cattaggcct	aggacacata	aataatttcc	ccagggccac	1860
tcagctagta	atcacagaac	tggagctcaa	atcccttgca	aaaaaaaaaa	aaaaaaaaaa	1920
aaaaaa						1926

<210> 116

<211> 1063

<212> DNA

<213> Homo sapiens

<400> 116

ccacgcgtcc	gaggagaaca	agtggactcg	ggtagcttct	atgagtacca	gaagactagg	60
tgtggctgtg	gctgtgttag	gagggttctt	atatgctgta	ggtggctctg	acgggacatc	120
tcctctcaac	acagtgggaa	gttacaatcc	tcaggaaaac	agatggcaca	ctatagcccc	180
tatggggacc	cggaggaaac	acctaggctg	tgtagtatat	caggacatga	tctatgctgt	240
aggaggtaga	gatgacacta	cagagctgag	cagtgtctgag	agatacaacc	ccagaaccaa	300
ccagtggctt	ccagtgggtg	ccatgacatc	acgccgtagt	ggagtgggcc	tggcagtggg	360
caatggacag	ctcatggcag	taggaggttt	tgatggcaca	acatacttga	agaccataga	420
agtttttgat	cctgatgccca	atacatggag	gttatatggc	gggatgaatt	accgtcggct	480
aggggggtgg	gtaggagtta	ttaaaatgac	acattgtgaa	tcccatattt	ggtgaacaca	540
gagaagacag	tcttgtatat	attcctctgt	attctgggga	gctttgacct	tggagctttg	600
tcagacttga	gaaaacatta	gaacaaattt	tattatttgc	cgtgtccctca	acaaatggaa	660
atacaatcca	atgaaagtac	ttcacctgca	agatgcacaa	taattttcaa	ctctgtgcag	720
aagaatattt	atttttggtt	ttaatttatc	atggtttttt	gttgttttct	ttttgaactt	780
atccttcctc	ccacaaaaaa	agaaaagaag	aaaaaattcc	aagcagcaaa	acttactttg	840
tttgtaaggt	attcatttag	gtttgaaaat	actatttaat	aaggggcagaa	gggctatata	900
tgatttggtc	attatttcta	gacactccat	ccacatgatt	ccactaacia	ggattaccag	960
gaataaaggt	aggtatgcaa	aatgtattag	ctaccattta	ttctcgcact	aaccaccaga	1020
agacttgaaa	atcttaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaa		1063

<210> 117

<211> 1615

<212> DNA

<213> Homo sapiens

<400> 117

ggcacgagct	cgtgccgctt	aagtttacag	gttcagatag	cttttctcac	accataggac	60
tcgtctatca	tttttgtgta	gtttttcttt	tttaatatgg	cttgtcttat	cagatttcct	120
gctattgggt	ccctccctta	ttccacctgg	cccttctttt	tctttatctt	tttatttttc	180
tcctgtttaa	cttttattcc	attttctcca	cttttctttt	tctgtgagcc	ataccctaga	240
aaagaacctt	agtggggccag	agttgagatg	caaattctta	gactactcta	gccccttgaa	300
ctcactccat	attggcacaag	ccagaaaatc	tgactgtttt	acttgctgtt	ctcaaatcaa	360
tccatcttac	tttccaggga	atatcttctg	tctgcttttg	atatgattcc	agaagccatg	420
tcacttctag	taaattactc	tgcttctcct	gtacactttc	aaatttctta	gggcctcttg	480
catataaagt	ggaaggtctc	atctctaaat	ttctagtaga	atccaactaa	aaacacatac	540
cctgagctga	gaccttctc	tgagcaaagg	aactcaccga	gtcactcttt	ggaactttca	600
agtcgtgctc	tttgtacgta	tcatatatac	tatgaacttg	ttttctctc	attgaaagat	660

aagatgtcag	ctttgcatgt	ttcctttatt	ccagtggaga	tccttcaggg	tttggtgagt	720
ggaaatctgg	gaggcacact	tgggccgact	gtcagcagcc	ccattgagca	agatgtgggtc	780
agtcccgttt	ccctgccccg	gagagcaaag	acctttggag	gatttgacag	ccgcttcagc	840
aagggtggtga	ctattctggt	aactatgggt	acaataatga	caaccaggaa	ttttatcagg	900
atacttatgg	gcaacagtgg	aagtagacaa	gtaagggcct	gaaaatgata	ctggcaagat	960
acgattggct	ctagatctac	attcttcaaa	aaaaaaaatt	ggcttaactg	tttcatcttt	1020
aagtacattt	tgctgccatt	tgtattgggc	tgaagaaatc	actattgtgt	atatactcaa	1080
gtctttttat	ttttcctctt	ttcataaatg	ctcttggaca	ttattgggct	tgcagagttc	1140
ccttattctg	gggattacaa	tgcttttatc	gtttcaggct	tcatttttagc	ttcaaaacaa	1200
gctgggcaca	ctgttaaatac	atgattttgc	agaacctttg	gttttggaca	gtttcatctt	1260
tttggatttg	ggatagatta	cataggagta	tggagtatgc	tgtaaataaa	aatacaagct	1320
agtgcctttgt	cttagtagtt	ttaagaaatt	aaagcaaaca	aatttaagtt	ttcttgtatt	1380
gaaaataacc	tatgattgta	tgttttgcac	tcctagaagt	aggttaactg	tgtttttaaa	1440
ttgttataac	ttcacacctt	tttgaaatct	gccctacaaa	atttgtttgg	cttaaactgc	1500
aaaagccgtg	acaatttggt	ctttgatgtg	attgtatttc	caatttcctg	ttcatgttaa	1560
gatttcaata	aaactaaaaa	atctattcaa	aacattaaaa	aaaaaaaaaa	aaaaa	1615

<210> 118

<211> 1221

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (697)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (700)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (701)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (712)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (720)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (722)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (742)

<223> n equals a,t,g, or c

<400> 118
 cactagtggga tccaaagaat tcggcacgag gagagacttg catagatcag ggagtatgtg 60
 aaaataaatg tcagttgcca gagaaagaat agatcaaaca gtttatttga atttctgatg 120
 agtttgggtga ggaagtgtat agcacttaac atttgggaagg cagacaaaag gttttgttgt 180
 tgggtgttatg ttctttttga attttagata cataatgcga ttttttttct ggccaatgct 240
 ccaggcaaaa ttgatgtctt tccactttct aaaacccatc atatttatga actctctgat 300
 actctgtctg aaacagtcgt gctcctgtga ggttgaaata tctctcctgc ctctatcaca 360
 gcagacgcac agaactgatc tgggattttc tctctcagga tcccaaaatg agccctttct 420
 caatcttgac aaacgtgcag cagaagccca ttgtgcagtt atgggttctgt gcctgctggg 480
 cagggattta aaggccagga gaagcagga aggccctgct ctctgctcca gccaggtggt 540
 aatatgcatt ctcaaactag ctaggaaaag gttctagatt cctttgaatc tattgctgaa 600
 tcccactgat tagcttttcc cactgacaga cattcaaatc tttgtttttg gagcttcctc 660
 tgcagacttg gtgtaaattt caaaaaaaga aaacaangcn ntgaagctta anacaggcan 720
 anacagttgg ctctctacct angcatgctt caaactcact atatgacttg gtaaatctctg 780
 ttactttaat gcatgcttgt tttcacatgt gtaaaatgag agagtatcag actagatcat 840
 ggatttgggtt cccctctttt tccccttctc tctogaagag tagaaccctt taaattaaat 900
 cttttataaa agttaagtgg ccgggagctg tggcccatgc ctgtaatccc agcactttag 960
 gagggcgaag cgggtggatc acctgaggtc gggagttcaa gaccagcctg accaactagg 1020
 agaaaccctg tctttactaa aaatacaaaa aattaggcag gcatggcggt gcatgcctgt 1080
 aatcccgact actcgggagg ctgaggtagg agaatcgctt gaacctggga ggcagagggt 1140
 gcggtgagcc gagatcatgt ccattgcact ccagcctggg caacaagaac aaaaatccgt 1200
 ctcaaaaaaa aaaaaaaaaa a 1221

<210> 119

<211> 1149

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (1120)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (1140)

<223> n equals a,t,g, or c

<400> 119
 gggtttctgg ggattggatt caaagagcag aagtatgatg ccaggggctg gtaaggactc 60
 agccatatag agatttcagc tgcattataa tgtatatttg agaaattgag ttcttattat 120
 cagaaaaattg acctatgtaa tgaaatacta gaagaaaact ctaaaattaga atggccttgt 180
 tatgaatttt taatgataag tactaacact taaaaggga aaaatttttc aaagaaaatg 240
 cctaaaaatg tgttgtatct gagatcatca tatatgtagt attacagtgc aaacactatt 300
 aatgttaatg aggtaggtgg atttttttaa tgaaaaaact tgaaatatta tcaggttgct 360
 cccatttatg gaatttaagc ttgtcagaaa gatccagata gctatattaa tattttatct 420
 gtatttgggt gcggttgctt ttaaaaaataa gttttcttat aagtcatttc aattttttgg 480
 tttagagtcc atatttcaaa ataaaaagtt aaaaaaagag taccttatgt aatttagcaa 540
 ggtattccat aaactcaggg aggaaaaaaa aaaaactaag atgagaggct agttataatg 600
 atctttatct atatatgca ggaaatgatt tcctctctca caacttagca taatttactg 660
 ctaatatttt tagccatggg ttgaaaaaaa ttaartggtt ttgtaagtgc taatcctttg 720
 ttactcattt aaactggtaa aatgtagtcc ccgctccctg aggaggacct gtccaaactc 780
 ttcaaacacc cacagccgcc tgccaggatg gactcgctgc tcattgcagg ccagataaac 840
 acttactgcc agaacatcaa ggagttcact gcccaaaact taggcaagct cttcatggcc 900
 caggctcttc aagaatacaa caactaagaa aaggaaagtt ccagaaaaga agttaacatg 960

aactcttgaa gtcacaccag ggcaactctt ggaagaaata tatttgcata ttgaaaagca	1020
cagaggattt cttagtggtc attgccgatt ttggctataa cagtgtcttt cttagccataa	1080
taaaataaaa caaatcttg aaaaaaaaaa aaaaaaaaaa gggsggccgc tctaagggn	1140
tccaagctt	1149

<210> 120

<211> 1515

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (69)

<223> n equals a,t,g, or c

<400> 120

aattcggcac gagggaaatt caagcacttt tcctaaaaga agggggaatg gatgctgaaa	60
caacacgtnt cccacaaagg gagcagacac tgggcttggt aagctgcccc ataccttccc	120
cacagaactg ggggtccggcc tccctgacat gcagatttcc acccagaaga cagagaagga	180
gccagtgggtc atggaatggg ctgggggtcaa agactgggtg cctgggagct gaggcagcca	240
ccgttttcagc ctggccagcc ctctggaccc cgaggttggc ccctactgtg acacacctac	300
catgcggaca ctcttcaacc tcctctgggt tgcctggcc tgacagccctg ttcacactac	360
cctgtcaaaag tcagatgcca aaaaagccgc ctcaaagacg ctgctggaga agagtcagtt	420
ttcagataag ccggtgcaag accgggggtt ggtgggtgacg gacctcaaag ctgagagtgt	480
ggttcttgag catcgagct actgctcggc aaaggcccg gacagacact ttgctgggga	540
tgtactgggc tatgtcactc catggaacag ccatggctac gatgtcacca aggtctttgg	600
gagcaagttc acacagatct caccgtctg gctgcagctg aagagacgtg gccgtgagat	660
gtttgaggtc acgggcctcc acgacgtgga ccaagggtgg atgagagctg tcaggaagca	720
tgccaagggc ctgcacatag tgcctcggct cctgtttgag gactggactt acgatgattt	780
ccggaacgtc tttagacagt aggatgagat agaggagctg agcaagaccg tgggtccaggt	840
ggcaaagaac cagcatttcg atggcttcgt ggtggagggtc tggaaaccagc tgctaagcca	900
gaagcgcgtg accgaccagc tgggcatgtt cagcacaag gattttgagc agctggcccc	960
cgtgctggat ggtttcagcc tcattgaccta cgactactct acagcgcatc agcctggccc	1020
taatgcaccc ctgtcctggg ttcgagcctg cgtccagggtc ctggaccgga agtccaagtg	1080
gcgaagcaaa atcctcctgg gggtcaactt ctatggtatg gactacgcga cctccaagga	1140
tgcccgtgag cctgttgtcg gggccaggta catccagaca ctgaaggacc acaggccccg	1200
yatgggtgtg gacagccagg yctcagagca ctctctcgag tacaagaaga gccgcagtgg	1260
gagggacgtc gtcttctacc caaccctgaa gtcctgagc gtgcccgtgg agctggcccc	1320
ggagctgggc gttggggtct ctatctggga gctgggcccag ggcctggact acttctacga	1380
cctgctctag gtgggcattg cggcctccgc ggtggacgtg ttcttttcta agccatggag	1440
tgagttagca ggtgtgaaat acaggccttc actccgttaa aaaaaaaaaa aaaaaaaaaa	1500
aaaaaaaaaa aaaaa	1515

<210> 121

<211> 1025

<212> DNA

<213> Homo sapiens

<400> 121

ggagaattgt tactggtgaa ttggttttca ggttttgagc atacatatac aatgtgtttc	60
aatgtctgct tgtaaaatta aaccatcatc tagaatagag ttaatttatt ataataagc	120
tacaaatagg tttttctaaa ccagagatgc actgcccttg tctaaagttc ttattgcact	180
ggttttatatg tgtagtggtg tttttattgtg tgttttttta atttgtaagt attctaagag	240
tttcttaata ctaaggttaa aattttcatg ttgacctgag ctttttgcaa atttgctttg	300
gctctattga tttgtccatt atgtgttagg caaatataac ttaagtggag ggggaagttt	360

atgaatataa	tatagctctg	tgttttaaac	ctcagaaaca	gatttgagtg	tttcagtatt	420
atagaaacag	tgatgactat	tcatgctctg	ctagtctatg	cctgcaactc	caaatgtttg	480
tggttcagta	tttccacact	acattttctgt	ttggtgacat	tgctcatttt	aacaaatag	540
accgagtcta	gtttttcttt	aaaaggatag	tttatgagta	atcttttaaa	ccatttccat	600
accatctgta	tataaccatt	tcggtagaga	acacactaca	ctgaaccctg	ctttagagct	660
gtgtgttgag	ctaaaaatat	aattttttta	aaattgacta	gcaaaatcta	tggccacact	720
gagaagcctt	tgaaaatggc	aaatactttt	catcaccaat	tgcccaattc	atctttcttc	780
tgcttctca	gccttgtagc	aaaggctaca	cagcagccca	cagtccacag	tctttttggg	840
aaaattggcc	tgccaccttc	tttaagctca	gtttattttt	gacttacttt	ctttgctgta	900
gttatgaacc	ttggggcatt	aaaatcccat	ggcaaggagc	ataagagatg	ttctcgtagc	960
tctgcgttgt	gtgaaatgtc	catcttagtt	ttgttaaaaa	aaaaaaaaaa	aaaaaaaaac	1020
tcgag						1025

<210> 122

<211> 2207

<212> DNA

<213> Homo sapiens

<400> 122

ggcacgagca	cgaatcagct	gcagggtctct	gttttgaaaa	agcagagata	cagaggcgaga	60
ggaaaaggg	ggactccctat	gtgacctgtt	cttagagcaa	gacaatcacc	atctgaattc	120
cagaagccct	gttcatgggt	ggggatattt	tctcgactgc	atggaatcag	aaagaagcaa	180
aaggatggga	aatgcctgca	ttcccctgaa	aagaattgct	tatttccctat	gtctcttctc	240
tgcgcttttg	ctgactgagg	ggaagaaaacc	agcgaaccaa	aatgccctgc	cgtgtgtact	300
tgtaccaaag	ataatgcttt	atgtgagaat	gccagatcca	ttccacgcac	cgttccctct	360
gatgttatct	cattatcctt	tgtgagatct	ggttttactg	aaatctcaga	agggagtttt	420
ttattcacgc	catcgctgca	gctcttgcta	ttcacatcga	actcctttga	tgtgatcagt	480
gatgatgctt	ttattggctt	ttcacatcta	gagtatttat	tcatagaaaa	caacaacatc	540
aagtcaattt	caagacatac	tttccgggga	ctaaagtcac	taattcactt	gagccttgca	600
aacaacaatc	tccagacact	cccaaaagat	attttcaaa	gcctggattc	tttaacaaat	660
gtggacctga	ggggtaattc	atttaattgt	gactgtaaac	tgaaatggct	agtggaaatgg	720
cttggccaca	ccaatgcaac	tggtgaagac	atctactgcy	aaggccccc	agaatacaag	780
aagcgcaaaa	tcaatagtct	ctcctcgaag	gatttcgatt	gcatcattac	agaatttgca	840
aagtctcaag	acctgcctta	tcaatcattg	ttcatagaca	ctttttctta	tttgaatgat	900
gagtatgtag	tcacgctca	gccttttact	ggaaaatgca	ttttccttga	atgggaccat	960
gtggaaaaga	ccttccggaa	ttatgacaac	attacaggca	catccactgt	agratgcaag	1020
cctatagtca	cttgaactca	gctctatgtt	attgtggccc	agctgtttgg	tggtctctac	1080
atctataagc	gagacagttt	tgcaaaataa	ttcataaaaa	tccaggatat	tgaaattctc	1140
aaaatccgaa	aaccatga	cattgaaaca	ttcaagattg	aaaacaactg	gtactttgtt	1200
gttgctgaca	gttcaaaagc	tggttttact	accatttaca	aatggaacgg	aaacggattc	1260
tactcccac	aatccttaca	cgcgtggtag	agggacactg	atgtggaata	tctagaaata	1320
gtcagaacac	ctcagacact	cagaacgcct	catttaattc	tgtctagtag	ttcccaacgt	1380
cctgtaattt	atcagtggaa	caaagcaaca	caattattca	ctaaccacac	tgacattctt	1440
aacatggagg	atgtgtacgc	agtgaagcac	ttctcagtga	aaggggacgt	gtacatttgc	1500
ttgacaagat	tcattgggtga	ttccaaagtc	atgaaatggg	gaggctcctc	gttccaggat	1560
attcagagga	tgccatcgcg	aggatccatg	gtgttccagc	ctcttcaaat	aaataattac	1620
caatatgcaa	ttcttggaag	tgattactcc	tttactcaag	tgtataactg	ggatgcagag	1680
aaagccaaat	ttgtgaaatt	tcaggaatta	aatgttcagg	caccaagatc	atccacacat	1740
gtgtccatta	ataagcgtaa	ttttcttttt	gcttccagtt	ttaagggaaa	tacacagatt	1800
tacaaacatg	tcatagttga	cttaagcgca	tgagacacca	aattctgtgg	ctgccaatcag	1860
aaattttcta	cagtacatga	cccggatgaa	ctcaatgcat	gatgactctt	cttatcacac	1920
ttgcaaatga	atgcctttca	aacattgaga	ctgctagaac	caagcactac	cagtatctcc	1980
atccttaact	gtccagtcca	gtgatgtggg	aagttacctt	ttataagaca	aaatttaatt	2040
gtgtaactgt	tccttgcaag	gaagatgtgt	aaataagcgt	ttaatgggat	ctgttactcc	2100
aaaaagaaat	attaatatgt	acttttccat	ttattttatc	atgtgtacag	aaacaactgc	2160
caaataaaat	gtttacattt	tctttcataa	aaaaaaaaaa	aaaaaa		2207

<210> 123
 <211> 1770
 <212> DNA
 <213> Homo sapiens

<400> 123
 gctgagtgtg agctgagcct gccccaccac caagatgata ctgagcttgc tgttcagcct 60
 tggggggccc ctgggctggg ggctgctggg ggcatggggc caggcttcca gtactagcct 120
 ctctgatctg cagagctcca ggacacctgg ggtctggaag gcagaggctg aggacaccag 180
 caaggacccc gttggacgta actgggtgcc ctacccaatg tccaagctgg tcaccttact 240
 agctctttgc aaaacagaga aattcctcat ccactcgcag cagccgtgtc cgcaggagct 300
 ccagactgcc agaaagtcaa agtcatgtac cgcattggccc acaagccagt gtaccaggctc 360
 aagcagaagg tgctgacctc tttggcctgg aggtgctgcc ctggctacac gggccccaac 420
 tgccagcacc acgattccat ggcaatccct gagcctgcag atcctgggtga cagccaccag 480
 gaacctcagg atggaccagt cagcttcaaa cctggccacc ttgctgcagt gatcaatgag 540
 gttgaggtgc aacaggaaca gcaggaacat ctgctgggag atctccagaa tgatgtgcac 600
 cgggtggcag acagcctgcc aggcctgtgg aaagccctgc ctggttaacct cacagctgca 660
 gtgatggaag caaatcaaac agggcacgaa gttccctgat agatccttgg agcaggtgct 720
 gctacccccc gtggacacct tctacaagt gcatttcagc cccatctgga ggagctttaa 780
 ccaaagcctg cacagcctta cccaggccat aagaaacctg tctcttgacg tggaggccaa 840
 ccgcccaggc atctccagag tccaggacag tgccgtggcc agggctgact tccaggagct 900
 tgggtgcaaa tttgaggcca aggtccagga gaacactcag agagtgggtc agctgcgaca 960
 ggacgtggag gaacgcctgc acgcccagca ctttaccctg caccgctcga tctcagagct 1020
 ccaagccgat gtggacacca aattgaagag gctgcacaag gctcaggagg ccccagggac 1080
 caatggcagt ctggtgttgg caacgcctgg ggctggggca aggcctgagc cggacagcct 1140
 gcaggccagg ctgggccagc tgcagaggaa cctctcagag ctgcacatga ccacggccccg 1200
 caggaggagg gagttgcagt acaccctgga ggacatgagg gccaccctga cccggcacgt 1260
 ggatgagatc aaggaaactgt actccgaatc ggacgagact ttcgatcaga ttagcaagg 1320
 ggagcggcag gtggaggagc tgcaggtgaa ccacacggcg ctccgtgagc tgcgcgtgat 1380
 cctgatggag aagtctctga tcatggagga gaacaaggag gaggtggagc ggcagctcct 1440
 ggagctcaac ctacgcctgc agcacctgca ggggtggccat gccgacctca tcaagtacgt 1500
 gaaggactgc aattgccaga agctctatct agacctggac gtcattccggg agggccagag 1560
 ggacgccacg cgtgccctgg aggagaccca ggtgagcctg gacgagcggc ggcagctgga 1620
 cggctcctcc ctgcaggccc tgcagaacgc cgtggacgcc gtgtcgttgg ccgtggacgc 1680
 gcacaaagcg gagggcgagc gggcgcgggc ggccacgtcg cggctccgga gccaaagtga 1740
 ggcgctggat yacgaggtgg ggcgctgaa 1770

<210> 124
 <211> 1034
 <212> DNA
 <213> Homo sapiens

<400> 124
 ggcacgagga aagtacgagt cggctcagcc tggagggacc caaccagagc ctggcctggg 60
 agccaggatg gccatccaca aagccttggt gatgtgcctg ggaactgcctc tcttctgtt 120
 cccaggggcc tgggcccagg gccatgtccc acccggtgc agccaaggcc tcaacccct 180
 gtactacaac ctgtgtgacc gctctggggc gtggggcatc gtccctggagg ccgtggctgg 240
 ggcgggcatt gtcaccacgt ttgtgtcac catcatcctg gtggccagcc tcccctttgt 300
 gcaggacacc aagaaacgga gcctgctggg gacccagcta agaggccggt gtcaccatac 360
 agcgggtaca atgggcagct gctgaccagt gtgtaccagc ccaactgagat ggccctgatg 420
 cacaagttc cgtccgaagg agcttacgac atcatcctcc caccggccac cgccaacagc 480
 cagggtgatg gcagtgccaa ctcgaccctg cgggctgaag acatgtactc ggccagagc 540
 caccaggcgg ccacaccgcc gaaagacggc aagaactctc aggtctttag aaacccctac 600
 gtgtgggact gagtcagcgg tggcgaggag aggcggtcgg atttggggag ggccctgagg 660

acctggcccc	gggcaagggg	ctctccaggc	tcctcctccc	cctggcaggc	ccagcaacat	720
gtgccccaga	tgtggaaggg	cctccctctc	tgccagtgtt	tgggtgggtg	tcattgggtgt	780
ccccaccac	tcctcagtgt	ttgtggagtc	gaggagccaa	ccccagcctc	ctgccaggat	840
cacctggcg	gtcacactcc	agccaaatag	tgttctcggg	gtggtggctg	ggcagcgcct	900
atgtttctct	ggagattcct	gcaacctcaa	gagacttccc	aggcgctcag	gcctggatct	960
tgctcctctg	tgaggaacaa	gggtgcctaa	taaatacatt	tctgctttat	taaaaaaaaa	1020
aaaaaaaaaa	aaaa					1034

<210> 125
 <211> 353
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (353)
 <223> Xaa equals stop translation

<400> 125

Met	Leu	Cys	Arg	Leu	Cys	Trp	Leu	Val	Ser	Tyr	Ser	Leu	Ala	Val	Leu	1	5	10	15
Leu	Leu	Gly	Cys	Leu	Leu	Phe	Leu	Arg	Lys	Ala	Ala	Lys	Pro	Ala	Glu	20	25	30	
Thr	Pro	Arg	Pro	Thr	Ser	Leu	Ser	Gly	Ala	Pro	Pro	Thr	Pro	Arg	His	35	40	45	
Ser	Arg	Cys	Pro	Pro	Asn	His	Thr	Val	Ser	Ser	Ala	Ser	Leu	Ser	Leu	50	55	60	
Pro	Ser	Arg	His	Arg	Leu	Phe	Leu	Thr	Tyr	Arg	His	Cys	Arg	Asn	Phe	65	70	75	80
Ser	Ile	Leu	Leu	Glu	Pro	Ser	Gly	Cys	Ser	Lys	Asp	Thr	Phe	Leu	Leu	85	90	95	
Leu	Ala	Ile	Lys	Ser	Gln	Pro	Gly	His	Val	Glu	Arg	Arg	Ala	Ala	Ile	100	105	110	
Arg	Ser	Thr	Trp	Gly	Arg	Trp	Gly	Asp	Gly	Leu	Gly	Pro	Ala	Leu	Lys	115	120	125	
Leu	Val	Phe	Leu	Leu	Gly	Val	Ala	Gly	Ser	Ala	Pro	Pro	Ala	Gln	Leu	130	135	140	
Leu	Ala	Tyr	Glu	Ser	Arg	Glu	Phe	Asp	Asp	Ile	Leu	Gln	Trp	Asp	Phe	145	150	155	160
Thr	Glu	Asp	Phe	Phe	Asn	Leu	Thr	Leu	Lys	Glu	Leu	His	Leu	Gln	Arg	165	170	175	
Trp	Val	Val	Ala	Ala	Cys	Pro	Gln	Ala	His	Phe	Met	Leu	Lys	Gly	Asp	180	185	190	
Asp	Asp	Val	Phe	Val	His	Val	Pro	Asn	Val	Leu	Glu	Phe	Leu	Asp	Gly				

Xaa

```
<220>
<221> SITE
<222> (108)
<223> Xaa equals any of the naturally occurring L-amino acids
```

```
<220>
<221> SITE
<222> (156)
<223> Xaa equals any of the naturally occurring L-amino acids
```

```
<220>  
<221> SITE  
<222> (158)  
<223> Xaa equals stop translation
```

<400> 126
Met Ser Trp Val Gly Leu Gly Arg Arg Gly His Leu Leu Leu Leu Ile
1 5 10 15

72

Asn Pro Arg Ala Leu Ala Gly Ile Arg Leu Pro Ser Pro Thr Gly Ala
 20 25 30
 Pro Ala Pro Gly Pro Cys Pro Pro Leu Cys Thr Pro His Cys Ser Arg
 35 40 45
 Glu His Pro Ala Gly Gly Thr Gly His Pro Ala Gly Val Trp Trp Arg
 50 55 60
 Arg Gly Cys Tyr Gly Gly Ser Cys Pro Met Gly Pro Val Arg Gly Ile
 65 70 75 80
 Leu Gly Gly Leu Pro Cys Arg Glu Glu Ala Leu Arg Arg His His Ser
 85 90 95
 Lys Pro Cys Trp Arg Pro Gly Gly Gln Ala Arg Xaa Leu Gly Ser Trp
 100 105 110
 Pro Leu Thr Ala Gly Arg Glu Pro Pro Arg Thr Ala Ser Thr Ala Pro
 115 120 125
 His Thr Ser Glu Pro Thr Ser Ser Phe Pro Arg Phe Pro Arg Ser Gln
 130 135 140
 Ala Trp Glu Asp Leu Pro Asp Ala Ala His His Xaa Ser Xaa
 145 150 155

<210> 127

<211> 554

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (39)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (199)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (201)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (202)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (228)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>
 <221> SITE
 <222> (420)
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>
 <221> SITE
 <222> (434)
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>
 <221> SITE
 <222> (440)
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>
 <221> SITE
 <222> (452)
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>
 <221> SITE
 <222> (554)
 <223> Xaa equals stop translation

<400> 127
 Met Lys Ile Ala Thr Val Ser Val Leu Leu Pro Leu Ala Leu Cys Leu
 1 5 10 15
 Ile Gln Asp Ala Ala Ser Lys Asn Glu Asp Gln Glu Met Cys His Glu
 20 25 30
 Phe Gln Ala Phe Met Lys Xaa Gly Lys Leu Phe Cys Pro Gln Asp Lys
 35 40 45
 Lys Phe Phe Gln Ser Leu Asp Gly Ile Met Phe Ile Asn Lys Cys Ala
 50 55 60
 Thr Cys Lys Met Ile Leu Glu Lys Glu Ala Lys Ser Gln Lys Arg Ala
 65 70 75 80
 Arg His Leu Ala Arg Ala Pro Lys Ala Thr Ala Pro Thr Glu Leu Asn
 85 90 95
 Cys Asp Asp Phe Lys Lys Gly Glu Arg Asp Gly Asp Phe Ile Cys Pro
 100 105 110
 Asp Tyr Tyr Glu Ala Val Cys Gly Thr Asp Gly Lys Thr Tyr Asp Asn
 115 120 125
 Arg Cys Ala Leu Cys Ala Glu Asn Ala Lys Thr Gly Ser Gln Ile Gly
 130 135 140
 Val Lys Ser Glu Gly Glu Cys Lys Ser Ser Asn Pro Glu Gln Asp Val
 145 150 155 160

Cys Ser Ala Phe Arg Pro Phe Val Arg Asp Gly Arg Leu Gly Cys Thr
 165 170 175
 Arg Glu Asn Asp Pro Val Leu Gly Pro Asp Gly Lys Thr His Gly Asn
 180 185 190
 Lys Cys Ala Met Cys Ala Xaa Leu Xaa Xaa Lys Glu Ala Glu Asn Ala
 195 200 205
 Lys Arg Glu Gly Glu Thr Arg Ile Arg Arg Asn Ala Glu Lys Asp Phe
 210 215 220
 Cys Lys Glu Xaa Glu Lys Gln Val Arg Asn Gly Arg Leu Phe Cys Thr
 225 230 235 240
 Arg Glu Ser Asp Pro Val Arg Gly Pro Asp Gly Arg Met His Gly Asn
 245 250 255
 Lys Cys Ala Leu Cys Ala Glu Ile Phe Lys Gln Arg Phe Ser Glu Glu
 260 265 270
 Asn Ser Lys Thr Asp Gln Asn Leu Gly Lys Ala Glu Glu Lys Thr Lys
 275 280 285
 Val Lys Arg Glu Ile Val Lys Leu Cys Ser Gln Tyr Gln Asn Gln Ala
 290 295 300
 Lys Asn Gly Ile Leu Phe Cys Thr Arg Glu Asn Asp Pro Ile Arg Gly
 305 310 315 320
 Pro Asp Gly Lys Met His Gly Asn Leu Cys Ser Met Cys Gln Ala Tyr
 325 330 335
 Phe Gln Ala Glu Asn Glu Glu Lys Lys Lys Ala Glu Ala Arg Ala Arg
 340 345 350
 Asn Lys Arg Glu Ser Gly Lys Ala Thr Ser Tyr Ala Glu Leu Cys Ser
 355 360 365
 Glu Tyr Arg Lys Leu Val Arg Asn Gly Lys Leu Ala Cys Thr Arg Glu
 370 375 380
 Asn Asn Pro Ile Gln Gly Pro Asp Gly Lys Val His Gly Asn Thr Cys
 385 390 395 400
 Ser Met Cys Glu Val Phe Phe Gln Ala Glu Glu Glu Glu Lys Lys Lys
 405 410 415
 Lys Glu Gly Xaa Ser Arg Asn Lys Arg Gln Ser Lys Ser Thr Ala Ser
 420 425 430
 Phe Xaa Glu Leu Cys Ser Glu Xaa Arg Lys Ser Arg Lys Asn Gly Arg
 435 440 445
 Leu Phe Cys Xaa Arg Glu Asn Asp Pro Ile Gln Gly Pro Asp Gly Lys
 450 455 460

75

Met His Gly Asn Thr Cys Ser Met Cys Glu Ala Phe Phe Gln Gln Glu
 465 470 475 480

Glu Arg Ala Arg Ala Lys Ala Lys Arg Glu Ala Ala Lys Glu Ile Cys
 485 490 495

Ser Glu Phe Arg Asp Gln Val Arg Asn Gly Thr Leu Ile Cys Thr Arg
 500 505 510

Glu His Asn Pro Val Arg Gly Pro Asp Gly Lys Met His Gly Asn Lys
 515 520 525

Cys Ala Met Cys Ala Ser Val Phe Lys Leu Glu Lys Lys Lys Lys Lys
 530 535 540

Lys Lys Lys Lys Lys Gly Arg Pro Leu Xaa
 545 550

<210> 128

<211> 308

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (308)

<223> Xaa equals stop translation

<400> 128

Met Asn Thr Val Leu Leu Ser Leu Leu Phe Ser Leu Pro Arg Ile Val
 1 5 10 15

Tyr Ala Met Ala Ala Asp Gly Leu Phe Phe Gln Val Phe Ala His Val
 20 25 30

His Pro Arg Thr Gln Val Pro Val Ala Gly Thr Leu Ala Phe Gly Leu
 35 40 45

Leu Thr Ala Phe Leu Ala Leu Leu Leu Asp Leu Glu Ser Leu Val Gln
 50 55 60

Phe Leu Ser Leu Gly Thr Leu Leu Ala Tyr Thr Phe Val Ala Thr Ser
 65 70 75 80

Ile Ile Val Leu Arg Phe Gln Lys Ser Ser Pro Pro Ser Ser Pro Gly
 85 90 95

Pro Ala Ser Pro Gly Pro Leu Thr Lys Gln Gln Ser Ser Phe Ser Asp
 100 105 110

His Leu Gln Leu Val Gly Thr Val His Ala Ser Val Pro Glu Pro Gly
 115 120 125

Glu Leu Lys Pro Ala Leu Arg Pro Tyr Leu Gly Phe Leu Asp Gly Tyr
 130 135 140

76

Ser Pro Gly Ala Val Val Thr Trp Ala Leu Gly Val Met Leu Ala Ser
 145 150 155 160
 Ala Ile Thr Ile Gly Cys Val Leu Val Phe Gly Asn Ser Thr Leu His
 165 170 175
 Leu Pro His Trp Gly Tyr Ile Leu Leu Leu Leu Thr Ser Val Met
 180 185 190
 Phe Leu Leu Ser Leu Leu Val Leu Gly Ala His Gln Gln Gln Tyr Arg
 195 200 205
 Glu Asp Leu Phe Gln Ile Pro Met Val Pro Leu Ile Pro Ala Leu Ser
 210 215 220
 Ile Val Leu Asn Ile Cys Leu Met Leu Lys Leu Ser Tyr Leu Thr Trp
 225 230 235 240
 Val Arg Phe Ser Ile Trp Leu Leu Met Gly Leu Ala Val Tyr Phe Gly
 245 250 255
 Tyr Gly Ile Arg His Ser Lys Glu Asn Gln Arg Glu Leu Pro Gly Leu
 260 265 270
 Asn Ser Thr His Tyr Val Val Phe Pro Arg Gly Ser Leu Glu Glu Thr
 275 280 285
 Val Gln Ala Met Gln Pro Pro Ser Gln Ala Pro Ala Gln Asp Pro Gly
 290 295 300
 His Met Glu Xaa
 305

<210> 129

<211> 167

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (167)

<223> Xaa equals stop translation

<400> 129

Met Ala Ala Ala Val Leu Ala Met Thr Leu Ala Pro Thr Val Ser Gly
 1 5 10 15
 Thr Thr Ser Lys Cys Ser Ser Arg Arg Trp Cys Pro Val Pro Ala Ser
 20 25 30
 Ser Ser Cys Val Ser His Leu Leu Gly Ser Gly Cys Ala Pro Cys Ala
 35 40 45
 Pro Trp Thr Ala His Pro Arg Gln Pro Ser Gln Cys Trp Ser Ala Arg
 50 55 60

77

Ala Pro Arg Arg Leu Gly Ser Arg Pro Arg Arg Tyr Leu Leu Thr Gly
 65 70 75 80

Gln Ala Asn Gly Ser Leu Ala Met Trp Asp Leu Thr Thr Ala Met Asp
 85 90 95

Gly Leu Gly Gln Ala Pro Ala Gly Gly Leu Thr Glu Gln Glu Leu Met
 100 105 110

Glu Gln Leu Glu His Cys Glu Leu Ala Pro Pro Ala Pro Phe Ser Ser
 115 120 125

Leu Met Gly Leu Ser Pro Gln Pro Leu Thr Pro His Leu Pro His Gln
 130 135 140

Pro Pro Leu Ser Leu Gln Gln His Leu Leu Val Trp Pro Pro Trp Glu
 145 150 155 160

Pro Lys Pro Pro Ala Gly Xaa
 165

<210> 130

<211> 306

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (306)

<223> Xaa equals stop translation

<400> 130

Met Ala Ala Gly Leu Ala Arg Leu Leu Leu Leu Gly Leu Ser Ala
 1 5 10 15

Gly Gly Pro Ala Pro Ala Gly Ala Ala Lys Met Lys Val Val Glu Glu
 20 25 30

Pro Asn Ala Phe Gly Val Asn Asn Pro Phe Leu Pro Gln Ala Ser Arg
 35 40 45

Leu Gln Ala Lys Arg Asp Pro Ser Pro Val Ser Gly Pro Val His Leu
 50 55 60

Phe Arg Leu Ser Gly Lys Cys Phe Ser Leu Val Glu Ser Thr Tyr Lys
 65 70 75 80

Tyr Glu Phe Cys Pro Phe His Asn Val Thr Gln His Glu Gln Thr Phe
 85 90 95

Arg Trp Asn Ala Tyr Ser Gly Ile Leu Gly Ile Trp His Glu Trp Glu
 100 105 110

Ile Ala Asn Asn Thr Phe Thr Gly Met Trp Met Arg Asp Gly Asp Ala
 115 120 125

78

Cys Arg Ser Arg Ser Arg Gln Ser Lys Val Glu Leu Ala Cys Gly Lys
 130 135 140
 Ser Asn Arg Leu Ala His Val Ser Glu Pro Ser Thr Cys Val Tyr Ala
 145 150 155 160
 Leu Thr Phe Glu Thr Pro Leu Val Cys His Pro His Ala Leu Leu Val
 165 170 175
 Tyr Pro Thr Leu Pro Glu Ala Leu Gln Arg Gln Trp Asp Gln Val Glu
 180 185 190
 Gln Asp Leu Ala Asp Glu Leu Ile Thr Pro Gln Gly His Glu Lys Leu
 195 200 205
 Leu Arg Thr Leu Phe Glu Asp Ala Gly Tyr Leu Lys Thr Pro Glu Glu
 210 215 220
 Asn Glu Pro Thr Gln Leu Glu Gly Gly Pro Asp Ser Leu Gly Phe Glu
 225 230 235 240
 Thr Leu Glu Asn Cys Arg Lys Ala His Lys Glu Leu Ser Lys Glu Ile
 245 250 255
 Lys Arg Leu Lys Gly Leu Leu Thr Gln His Gly Ile Pro Tyr Thr Arg
 260 265 270
 Pro Thr Glu Thr Ser Asn Leu Glu His Leu Gly His Glu Thr Pro Arg
 275 280 285
 Ala Lys Ser Pro Glu Gln Leu Arg Gly Asp Pro Gly Leu Arg Gly Ser
 290 295 300
 Leu Xaa
 305

<210> 131
 <211> 220
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (56)
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>
 <221> SITE
 <222> (58)
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>
 <221> SITE
 <222> (204)
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (209)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (220)

<223> Xaa equals stop translation

<400> 131

Met Pro Cys Leu Glu Ala Val Ala Leu Ile Leu Leu Ile Leu Leu Val
 1 5 10 15

Pro Asp Pro Pro Arg Gly Ala Ala Glu Thr Gln Gly Glu Gly Ala Val
 20 25 30

Gly Gly Phe Arg Ser Ser Trp Cys Glu Asp Val Arg Tyr Leu Gly Lys
 35 40 45

Asn Trp Ser Phe Val Trp Ser Xaa Leu Xaa Val Thr Ala Met Ala Phe
 50 55 60

Val Thr Gly Ala Leu Gly Phe Trp Ala Pro Lys Phe Leu Leu Glu Ala
 65 70 75 80

Arg Val Val His Gly Leu Gln Pro Pro Cys Phe Gln Glu Pro Cys Ser
 85 90 95

Asn Pro Asp Ser Leu Ile Phe Gly Ala Leu Thr Ile Met Thr Gly Val
 100 105 110

Ile Gly Val Ile Leu Gly Ala Glu Ala Ala Arg Arg Tyr Lys Lys Val
 115 120 125

Ile Pro Gly Ala Glu Pro Leu Ile Cys Ala Ser Ser Leu Leu Ala Thr
 130 135 140

Ala Pro Cys Leu Tyr Leu Ala Leu Val Leu Ala Pro Thr Thr Leu Leu
 145 150 155 160

Ala Ser Tyr Val Phe Leu Gly Leu Gly Glu Leu Leu Leu Ser Cys Asn
 165 170 175

Trp Ala Val Val Ala Asp Ile Leu Leu Ser Val Val Val Pro Arg Cys
 180 185 190

Arg Gly Thr Ala Glu Ala Leu Gln Ile Thr Val Xaa His Ile Leu Gly
 195 200 205

Xaa Leu Ala Ala Leu Ser His Arg Thr Tyr Leu Xaa
 210 215 220

<210> 132

<211> 99

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (99)

<223> Xaa equals stop translation

<400> 132

Met	Met	Asn	Gln	His	Leu	Leu	Glu	Ser	Phe	Gly	Ser	Pro	Ser	Ser	Leu
1				5					10					15	

Phe	Ile	Val	Phe	Ile	Leu	Leu	Ile	Trp	Met	Leu	Gln	Arg	Cys	Lys	Asp
		20					25						30		

Phe	Phe	Leu	Cys	Cys	Tyr	Arg	Val	Val	Leu	Thr	Pro	Ser	Phe	Trp	Gln
		35					40					45			

Lys	His	Gln	His	Pro	Asp	Pro	Lys	Ile	Lys	His	His	Leu	Lys	Leu	Tyr
	50					55						60			

Ser	Leu	Lys	Tyr	Ser	Ser	Ser	Gly	Gln	Asn	Asn	Phe	Arg	Lys	Asp	Lys
	65				70					75					80

His	Trp	Leu	Ser	Gly	His	Thr	Glu	Glu	Ala	Asn	Leu	Ile	Lys	Glu	Glu
				85					90					95	

Trp Lys Xaa

<210> 133

<211> 61

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (61)

<223> Xaa equals stop translation

<400> 133

Met	Thr	Ser	Ser	Leu	Phe	Ile	Phe	Leu	Phe	Leu	Trp	Phe	Cys	Pro	Pro
1				5					10					15	

Pro	Arg	Ile	Ser	Phe	Val	Leu	Cys	Trp	Pro	Gln	Pro	His	Ser	Gln	Val
		20						25					30		

His	Ile	Gln	His	Glu	Lys	Ala	Asp	His	Leu	Phe	Gln	Ser	Leu	Lys	Gln
		35					40					45			

Lys	Ala	Pro	Gly	Leu	Leu	Gln	Trp	Ala	Arg	Ile	Val	Xaa
	50					55					60	

<210> 134

<211> 248

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (14)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (141)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (248)

<223> Xaa equals stop translation

<400> 134

Met Ala Val Pro Ala Leu Thr Pro Ala Ala Val Arg Ala Xaa Gly Leu
1 5 10 15

Leu Gly Val Ser Trp Thr Trp Ala Leu Phe Thr Pro Leu Val Ala Leu
20 25 30

Gly Arg Glu Gly Gly Ser Gln Asp Ser Ala Thr Thr Pro Ser Arg Pro
35 40 45

Pro Gly Arg Pro Arg Ile Val Asp Ile Ala Thr Ile Val His Cys Tyr
50 55 60

Ala Glu Glu Arg Gln Ser Ala Glu Asp Tyr Glu Lys Glu Glu Ser His
65 70 75 80

Arg Gln Arg Arg Leu Lys Glu Arg Glu Arg Ile Gly Glu Leu Gly Ala
85 90 95

Pro Glu Val Trp Gly Pro Ser Pro Lys Phe Pro Gln Leu Asp Ser Asp
100 105 110

Glu His Thr Pro Val Glu Asp Glu Glu Glu Val Thr His Gln Lys Ser
115 120 125

Ser Ser Ser Asp Ser Asn Ser Glu Glu His Arg Lys Xaa Lys Thr Ser
130 135 140

Arg Ser Arg Asn Lys Lys Lys Arg Lys Asn Lys Ser Ser Lys Arg Lys
145 150 155 160

His Arg Lys Tyr Ser Asp Ser Asp Ser Asn Ser Glu Ser Asp Thr Asn
165 170 175

Ser Asp Ser Asp Asp Asp Lys Lys Arg Val Lys Ala Lys Lys Lys Lys
180 185 190

Lys Lys Lys Lys His Lys Thr Lys Lys Lys Lys Asn Lys Lys Thr Lys
195 200 205

82

Lys Glu Ser Ser Asp Ser Ser Cys Lys Asp Ser Glu Glu Asp Leu Ser
 210 215 220

Glu Ala Thr Trp Asp Gly Ala Ala Lys Cys Gly Arg Tyr Tyr Gly Phe
 225 230 235 240

Asn Arg Ala Arg Ser Thr Tyr Xaa
 245

<210> 135
 <211> 41
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (41)
 <223> Xaa equals stop translation

<400> 135
 Met Val Cys Phe Tyr Ala Leu Leu Leu Cys Phe Leu Ser Ser Val Glu
 1 5 10 15

Ile Gly Pro Leu Ser Trp Leu Leu Cys Leu Ser His Ile Lys Cys His
 20 25 30

Phe Thr Ala Leu Pro Phe Glu Ala Xaa
 35 40

<210> 136
 <211> 75
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (75)
 <223> Xaa equals stop translation

<400> 136
 Met Leu His Leu Phe Cys Ser Gln Pro Leu Gly Leu Leu Phe Leu Leu
 1 5 10 15

Ile Phe Leu Gly Leu Asp Ser Leu Pro Arg Cys Leu Thr Ala Thr Arg
 20 25 30

Leu Gln Ser Pro Ile Ile Ile Phe Ser Thr Leu Ser Cys Ile Cys Ser
 35 40 45

Thr Ser Trp Leu Glu Leu Cys Ser Val Tyr Phe Leu Thr Leu Asn Tyr
 50 55 60

Leu His Val Val Pro Pro Cys Phe Leu Ile Xaa
 65 70 75

<210> 137
 <211> 75
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (75)
 <223> Xaa equals stop translation

<400> 137
 Met Gly Val Leu Thr Arg Glu Leu Phe Gly Val Val Gly Met Leu Tyr
 1 5 10 15
 Ile Leu Ile Val Gly Met Val Thr Trp Leu Asp Ala Phe Val Lys Thr
 20 25 30
 His Leu Met Val Met Gln Asn Glu Tyr Ile Leu Phe Tyr Val Asn Tyr
 35 40 45
 Thr Ser Lys Leu Asn Phe Phe Lys Lys Phe Leu Leu Lys Ser Lys Asp
 50 55 60
 Ile Cys Gly Ala Ser Cys Lys Phe Tyr Cys Xaa
 65 70 75

<210> 138
 <211> 58
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (58)
 <223> Xaa equals stop translation

<400> 138
 Met Lys Val Leu Leu Ser Leu Ser Leu Val Gly Leu Phe Ile Gly Phe
 1 5 10 15
 Ser Asp Ala Val Leu Asn Glu Thr Cys Arg Phe Trp Ile Asn Thr Ser
 20 25 30
 Ser Lys Gly Asn Leu Gln Ile Leu Lys Asn Gln Ile Gln Ile Ile Asp
 35 40 45
 Arg Leu Arg Lys Met Pro Ala Ser Ala Xaa
 50 55

<210> 139
 <211> 173
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (76)
 <223> Xaa equals any of the naturally occurring L-amino acids

 <220>
 <221> SITE
 <222> (124)
 <223> Xaa equals any of the naturally occurring L-amino acids

 <400> 139
 Met Leu Gly Ser Pro Cys Leu Leu Trp Leu Leu Ala Val Thr Phe Leu
 1 5 10 15

 Val Pro Arg Ala Gln Pro Leu Ala Pro Gln Asp Phe Glu Glu Glu
 20 25 30

 Ala Asp Glu Thr Glu Thr Ala Trp Pro Pro Leu Pro Ala Val Pro Cys
 35 40 45

 Asp Tyr Asp His Cys Arg His Leu Gln Val Pro Cys Lys Glu Leu Gln
 50 55 60

 Arg Val Gly Pro Ala Ala Cys Leu Cys Pro Gly Xaa Ser Ser Pro Ala
 65 70 75 80

 Gln Pro Pro Asp Pro Pro Arg Met Gly Glu Val Arg Ile Ala Ala Glu
 85 90 95

 Glu Gly Arg Ala Val Val His Trp Cys Ala Pro Phe Ser Pro Val Leu
 100 105 110

 His Tyr Trp Leu Leu Leu Trp Asp Gly Ser Glu Xaa Arg Arg Arg Gly
 115 120 125

 Pro Pro Leu Asn Ala Thr Val Arg Arg Ala Glu Leu Lys Gly Leu Lys
 130 135 140

 Pro Gly Gly Ile Tyr Val Val Cys Val Val Ala Ala Asn Glu Ala Gly
 145 150 155 160

 Ala Ser Arg Val Pro Gln Ala Gly Gly Glu Gly Leu Glu
 165 170

<210> 140
 <211> 46
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (46)
 <223> Xaa equals stop translation

<400> 140
 Met Thr Ile His Ala Leu Leu Val Tyr Ala Cys Asn Ser Lys Cys Leu

1 5 85 10 15
 Trp Phe Ser Ile Ser His Leu His Phe Cys Leu Val Thr Leu Leu Ile
 20 25 30
 Leu Thr Asn Met Thr Glu Ser Ser Phe Ser Leu Lys Gly Xaa
 35 40 45

<210> 141
 <211> 58
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (58)
 <223> Xaa equals stop translation

<400> 141
 Met Val Tyr Arg Ala Phe Leu Ile Ile Ile Leu Arg Phe Ile Leu Ile
 1 5 10 15
 Phe Leu Phe Lys Leu Asn Tyr Ser Lys Leu Cys Pro Glu Ile Pro Phe
 20 25 30
 Gly Leu Lys Phe Phe Ser Phe Val Cys Ile Lys Val Gln Ile Lys Lys
 35 40 45
 Thr Ser Arg Lys Arg Arg Pro Tyr Leu Xaa
 50 55

<210> 142
 <211> 67
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (67)
 <223> Xaa equals stop translation

<400> 142
 Met Phe Val Glu Arg Trp Leu Pro Cys Phe Leu Val Val Ala Val Val
 1 5 10 15
 Val Trp Val Phe Ala Cys Gly Pro Val Glu Asp Lys Glu Asp Ser Phe
 20 25 30
 Gly Trp Ser Ser Tyr Phe Leu Ala Ser Gly Leu Pro Pro Leu Leu Phe
 35 40 45
 Glu Ala Ser Gln Thr Arg Thr Val Arg Ala Gly Arg Leu Gly Val Phe
 50 55 60
 Val Cys Xaa

65

<210> 143
 <211> 53
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (53)
 <223> Xaa equals stop translation

<400> 143
 Met Ile Phe Lys Leu Leu Ile Phe Arg Ile Phe Phe His Glu Leu Ala
 1 5 10 15
 Leu Ala Leu Cys Ile Ser Asn Leu Val Ser Leu Pro Trp Leu Ser Tyr
 20 25 30
 Phe Trp Cys Pro Glu Met Gln Asn Leu Phe Leu Leu Asp Thr His Ile
 35 40 45
 Trp Val Leu Met Xaa
 50

<210> 144
 <211> 66
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (66)
 <223> Xaa equals stop translation

<400> 144
 Met Val Leu Ser Val Ala Leu Leu His Ala Leu Ser His Leu Met Pro
 1 5 10 15
 Cys Lys Thr Cys Leu Ala Ser Thr Ser Pro Ser Ala Met Ile Val Ser
 20 25 30
 Phe Leu Arg Pro Pro Gln Pro Ala Met Trp Asn Cys Glu Ser Ile Lys
 35 40 45
 Pro Phe Leu Phe Ile His Tyr Pro Val Ser Gly Ser Ile Phe Ile Ala
 50 55 60
 Val Xaa
 65

<210> 145
 <211> 57
 <212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (57)

<223> Xaa equals stop translation

<400> 145

Met Val Ala Ile Leu Leu Arg Glu Leu Pro Leu Ala Phe Leu Leu Val
1 5 10 15

Gly Ser Ser Gly Asp Lys Phe Cys Phe Thr Ser Ser Glu Asn Val Leu
20 25 30

Leu Ser Phe Ser Phe Leu Lys Asp Ile Phe Ala Gly Tyr Lys Asn Ser
35 40 45

Gly Leu Met Val Leu Phe Ile Val Xaa
50 55

<210> 146

<211> 67

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (67)

<223> Xaa equals stop translation

<400> 146

Met Ser Asn Phe Ile Ser Ile Thr Cys Leu Val Phe Thr Ile Leu Gly
1 5 10 15

His Leu Val Ser Leu Gln Val Ala His Ser Ser Val Phe Glu Phe Lys
20 25 30

Thr Leu Tyr Val Leu Lys Thr Asn Arg Tyr Ser Gln Ser Leu Phe Arg
35 40 45

His Phe Cys His Leu Ser Phe Ile Arg Thr Arg Lys Ile Phe Leu Lys
50 55 60

Asn Asn Xaa
65

<210> 147

<211> 49

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (49)

<223> Xaa equals stop translation

<400> 147

Met Met Lys Tyr Phe Phe Asp Val Val Val Phe Leu Thr Phe Phe Leu
 1 5 10 15

Val Phe Ser Leu Ser Ile Phe Leu Ser Asp Glu Glu Phe Pro Val Ser
 20 25 30

Arg Thr Gln Asn Ile Gly Leu Cys His Phe Asn Pro Ser Phe Ser Glu
 35 40 45

Xaa

<210> 148

<211> 89

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (89)

<223> Xaa equals stop translation

<400> 148

Met Leu Leu Leu Cys Leu Tyr Cys Thr Phe Phe Leu Met Pro Phe Ile
 1 5 10 15

Ile Lys Tyr Thr Cys Phe His Leu Val Phe Gly Gln Ile Pro Val Thr
 20 25 30

Val His Val Asn Ile Trp Gln His Lys Asn Val Thr Phe Phe Ile Leu
 35 40 45

His Cys Gly Ile Pro Ala Leu Thr Arg Asp Ser Ala Ala Leu Thr Tyr
 50 55 60

Ser Asn Asp Gly Thr Val Ile Glu Thr Leu Leu Phe Leu Ile Leu Tyr
 65 70 75 80

Leu Asp Leu Asn Ile Ile Cys Cys Xaa
 85

<210> 149

<211> 77

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (77)

<223> Xaa equals stop translation

<400> 149

Met Thr Leu Tyr Ser Lys Leu Leu Trp Leu Phe Lys Gly Glu Leu Leu

PCT/US98/27059

```

<210> 150
<211> 42
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (42)
<223> Xaa equals stop translation

<400> 150
Met Asn Leu Ser Phe Leu Ser Phe Phe Leu Phe Phe Tyr Leu Leu Trp
 1             5             10             15
Ser Pro Ala Glu Ser Val Tyr Lys Lys Gly Met Val Lys Lys Asn Leu
          20             25             30
Ser His Ser Ile Val Glu Lys Ile Lys Xaa
      35             40

```

```

<210> 151
<211> 46
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (46)
<223> Xaa equals stop translation

<400> 151
Met Asn Ala Leu Pro Asn Leu Ala Trp Leu Pro Phe Val Pro Ala Leu
 1             5             10             15
Ala Ala Ala Ser Pro Ala Gly Leu Ala Ala Pro Glu Ser Arg Asp Val
      20             25             30
Pro Phe Pro Val Ser Pro Ala Thr Gln Leu Asn Ile Gly Xaa
 35             40             45

```

SUBSTITUTE SHEET (RULE 26)

<210> 152
 <211> 42
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (42)
 <223> Xaa equals stop translation

<400> 152
 Met Leu His Leu Leu Cys Leu Gly Leu His Leu Val Pro Pro Gly Leu
 1 5 10 15
 Leu Ser Val Asn Ser Leu Gln Ser Thr Gln Cys Ser Leu Phe Ser Ala
 20 25 30
 Ala Lys Phe Phe Ser Ile Val Gln Val Xaa
 35 40

<210> 153
 <211> 44
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (44)
 <223> Xaa equals stop translation

<400> 153
 Met Pro Tyr Met Phe Arg Pro Ala Phe Leu Asn Cys Gly Thr Phe Ala
 1 5 10 15
 Ile Phe Gly Gln Leu Asn Ser Val Val Gly Ala Val Leu Cys Ile Ala
 20 25 30
 Gly Cys Leu Ala Ala Ser Leu Ala Ser Thr Tyr Xaa
 35 40

<210> 154
 <211> 123
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (123)
 <223> Xaa equals stop translation

<400> 154
 Met Pro Pro Leu Ala Pro Gln Leu Cys Arg Ala Val Phe Leu Val Pro
 1 5 10 15
 Ile Leu Leu Leu Leu Gln Val Lys Pro Leu Asn Gly Ser Pro Gly Pro

20 91 30
 25
 Lys Asp Gly Ser Gln Thr Glu Lys Thr Pro Ser Ala Asp Gln Asn Gln
 35 40 45
 Glu Gln Phe Glu Glu His Phe Val Ala Ser Ser Val Gly Glu Met Trp
 50 55 60
 Gln Val Val Asp Met Ala Gln Gln Glu Glu Asp Gln Ser Ser Lys Thr
 65 70 75 80
 Ala Ala Val His Lys His Ser Phe His Leu Ser Phe Cys Phe Ser Leu
 85 90 95
 Ala Ser Val Met Val Phe Ser Gly Gly Pro Leu Arg Arg Thr Phe Pro
 100 105 110
 Asn Ile Gln Leu Cys Phe Met Leu Thr His Xaa
 115 120

<210> 155
 <211> 42
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (42)
 <223> Xaa equals stop translation

<400> 155
 Met Lys Gln Phe Gly Phe Gly His Pro Ile Lys Leu Leu Lys Thr Lys
 1 5 10 15
 Leu Cys Arg Ile Val Phe Tyr Leu Val Phe Phe Val Trp Pro Gln Ser
 20 25 30
 Ser Val Ile Arg Glu Ala Thr Gln Thr Xaa
 35 40

<210> 156
 <211> 56
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (56)
 <223> Xaa equals stop translation

<400> 156
 Met Val Leu Ala Ala Pro Leu Val Ala Phe Pro Cys Ile Leu Leu Phe
 1 5 10 15
 Ala Phe Ser Pro Ser Ala Val Arg Asp His Val Gly Asp Ser Arg Ser

92

20

25

30

Asp Val Pro Ile Phe Ala Cys Leu Ala Leu Ala Ser Leu Ala Leu Gly
 35 40 45

Ser Val Leu Leu Val Ala Phe Xaa
 50 55

<210> 157
 <211> 45
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (45)
 <223> Xaa equals stop translation

<400> 157
 Met Met Lys Met Val Leu Gly Leu Phe Phe Leu Met Asn Leu Leu Ser
 1 5 10 15

Gly Lys Lys Ser Val Arg His His Ser Lys Asn Tyr Val Lys Lys Met
 20 25 30

Gln Thr Phe Gln Phe Pro Arg Val Tyr Lys Leu Met Xaa
 35 40 45

<210> 158
 <211> 86
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (86)
 <223> Xaa equals stop translation

<400> 158
 Met Lys Lys Val Leu Leu Leu Ile Thr Ala Ile Leu Ala Val Ala Val
 1 5 10 15

Gly Phe Pro Val Ser Gln Asp Gln Glu Arg Glu Lys Arg Ser Ile Ser
 20 25 30

Asp Ser Asp Glu Leu Ala Ser Gly Phe Phe Val Phe Pro Tyr Pro Tyr
 35 40 45

Pro Phe Arg Pro Leu Pro Pro Ile Pro Phe Pro Arg Phe Pro Trp Phe
 50 55 60

Arg Arg Asn Phe Pro Ile Pro Ile Pro Glu Ser Ala Pro Thr Thr Pro
 65 70 75 80

Leu Pro Ser Glu Lys Xaa

<210> 159
 <211> 45
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (45)
 <223> Xaa equals stop translation

<400> 159
 Met Ile Cys Leu Cys Ser Ile Lys Met Leu Leu Leu Phe Cys Gln Leu
 1 5 10 15
 Thr Phe Ala Leu Ile Thr Cys Ile Asn Leu Gln Ser Leu Tyr Leu Phe
 20 25 30
 Ser Tyr Gln Gln Ile Ile Gly Ile His Ser His Val Xaa
 35 40 45

<210> 160
 <211> 69
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (69)
 <223> Xaa equals stop translation

<400> 160
 Met Trp Leu Arg Gly Ile His Pro Phe Leu Trp Leu Ser Gly Ile His
 1 5 10 15
 Ser Phe Pro Trp Leu Ser Gly Gly Pro Ser Leu Gly Thr Ser Ser Glu
 20 25 30
 Gln Pro Thr Ser Leu Glu Asp Gly Lys Leu Ile Cys Leu Phe Thr Asp
 35 40 45
 Phe Ser Gly Ser Ser Phe Gly Leu Phe Met Arg Glu Ala Ala Lys Asn
 50 55 60
 Ile Ser Gln Met Xaa
 65

<210> 161
 <211> 53
 <212> PRT
 <213> Homo sapiens

<220>

<221> SITE
 <222> (53)
 <223> Xaa equals stop translation

<400> 161
 Met Leu Tyr Asp Ser Asn Leu Cys Ser Val Trp His Leu Tyr Leu Ile
 1 5 10 15
 Leu His Leu Cys Lys Thr Phe Val Tyr Cys Gly Cys Val His Ser Ser
 20 25 30
 Tyr Leu Ile Ser Gly Thr Val Asn Thr Gln Tyr Phe Ile Val Gln Thr
 35 40 45
 Val Leu Leu Phe Xaa
 50

<210> 162
 <211> 57
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (57)
 <223> Xaa equals stop translation

<400> 162
 Met Arg Val Lys Ile Ser Tyr Leu Met Ile Ala Leu Thr Val Val Gly
 1 5 10 15
 Cys Ile Phe Met Val Ile Glu Gly Lys Lys Ala Ala Gln Arg His Glu
 20 25 30
 Thr Leu Thr Ser Leu Asn Leu Glu Lys Lys Ala Arg Leu Lys Glu Glu
 35 40 45
 Ala Ala Met Lys Ala Lys Thr Glu Xaa
 50 55

<210> 163
 <211> 56
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (56)
 <223> Xaa equals stop translation

<400> 163
 Met Arg Glu Lys Thr Gly Ala Leu Pro Arg Cys Leu Gly Leu Leu Gly
 1 5 10 15
 Val Gly Leu Leu Trp Arg Trp Cys Gly Arg Arg Ala Arg Ala Gly Val

95

20

25

30

Gly Lys Ala Trp Ser Ala Thr Arg Ser Pro Ser Asp Ser Cys Phe Pro
 35 40 45

Gly Val Ala Arg Val Gly Ile Xaa
 50 55

<210> 164
 <211> 48
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (48)
 <223> Xaa equals stop translation

<400> 164
 Met His Gly His Thr Ser Ser Leu Pro Pro Ser Leu Leu Ser Ser Leu
 1 5 10 15
 Pro Ser Gly Leu Leu Ala Leu Phe Val Phe Pro Phe Leu Ile Leu Leu
 20 25 30
 Leu His Ala Glu Asp Leu Pro Tyr Tyr Tyr Phe Gly Asn Ile Glu Xaa
 35 40 45

<210> 165
 <211> 130
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (130)
 <223> Xaa equals stop translation

<400> 165
 Met Ser Ala Ser Ser Leu His Arg Leu Pro Val Leu Met Ala Leu Phe
 1 5 10 15
 Pro Phe Gln Ala Ala Ala Gly Ser Leu Gly Leu Gln Pro Pro Pro
 20 25 30
 Thr Pro Met Lys Gly Lys Pro Ser Ile Met Leu Pro Pro Gln Tyr Lys
 35 40 45
 Arg Arg Glu Gly Leu Lys Lys Lys Lys Lys Ile Gln Lys Val Ala
 50 55 60
 Leu Val Ser Phe Gly Arg Ala Asp Ser Ile Val Gly Asp Gly Leu Pro

WO 99/31117

PCT/US98/27059

65 70 96 75 80
 Thr Asn Gln Gly Asp Lys Cys Gln Arg Glu Arg Thr Met Pro Gly Ser
 85 90 95
 Lys His Ile Ser Pro Gln Thr Pro Gln Val Gly Lys Gln Ala Arg Gly
 100 105 110
 Ser Thr Asn Pro Ser Gly Arg Pro Gly Val Gln Met Leu Tyr Ser Ser
 115 120 125
 Ile Xaa
 130

<210> 166
 <211> 105
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (105)
 <223> Xaa equals stop translation

<400> 166
 Met Leu Trp Leu Leu Phe Phe Leu Val Thr Ala Ile His Ala Glu Leu
 1 5 10 15
 Cys Gln Pro Gly Ala Glu Asn Ala Phe Lys Val Arg Leu Ser Ile Arg
 20 25 30
 Thr Ala Leu Gly Asp Lys Ala Tyr Ala Trp Asp Thr Asn Glu Glu Tyr
 35 40 45
 Leu Phe Lys Ala Met Val Ala Phe Ser Met Arg Lys Val Pro Asn Arg
 50 55 60
 Glu Ala Thr Glu Ile Ser His Val Leu Leu Cys Asn Val Thr Gln Arg
 65 70 75 80
 Tyr His Ser Gly Leu Trp Leu Gln Thr Leu Gln Lys Ile Thr Pro Phe
 85 90 95
 Leu Leu Leu Arg Cys Asn Gln Pro Xaa
 100 105

<210> 167
 <211> 77
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (77)
 <223> Xaa equals stop translation

SUBSTITUTE SHEET (RULE 26)

<400> 167

Met Thr Lys Ala Arg Leu Phe Arg Leu Trp Leu Val Leu Gly Ser Val
 1 5 10 15

Phe Met Ile Leu Leu Ile Ile Val Tyr Trp Asp Ser Ala Ala Pro Arg
 20 25 30

Thr Ser Thr Cys Thr Arg Pro Ser Leu Gly Arg Thr Arg Gly Arg Arg
 35 40 45

Cys Pro Arg Pro Gly Arg Thr Gly Gln Gly Ala His Gly Arg Leu Arg
 50 55 60

Cys Arg Arg Val Ser Gly Gln Phe Leu Met Leu Ala Xaa
 65 70 75

<210> 168

<211> 355

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (355)

<223> Xaa equals stop translation

<400> 168

Met Trp Arg Leu Trp Pro Gly Ser Pro Leu Val Pro Leu Ser Trp Leu
 1 5 10 15

Trp Pro Ala Arg Ala Ala Phe Leu Ser Gly Pro Trp Thr Leu Pro Pro
 20 25 30

Cys Leu Pro Asp Pro Leu Leu Ala Val Pro Lys Cys Cys Leu Thr Leu
 35 40 45

Gly Ile His Leu Leu Pro Ala Trp Pro Gly Pro Pro Val Gly Gly Gly
 50 55 60

Cys Ser Gln Leu His Arg Gly Cys Cys Tyr Pro Gly Met Gly Cys Leu
 65 70 75 80

Asn Arg Asp Leu Cys Pro Pro Ser Leu Val Ser Arg Arg Trp Gly Asp
 85 90 95

Gln Leu Leu Trp Ser Pro Asp Gly Ser Lys Ile Leu Ala Thr Thr Pro
 100 105 110

Ser Ala Val Phe Arg Val Trp Glu Ala Gln Met Trp Thr Cys Glu Arg
 115 120 125

Trp Pro Thr Leu Ser Gly Arg Cys Gln Thr Gly Cys Trp Ser Pro Asp
 130 135 140

Gly Ser Arg Leu Leu Phe Thr Val Leu Gly Glu Pro Leu Ile Tyr Ser

145 150 155 160
 Leu Ser Phe Pro Glu Arg Cys Gly Glu Gly Lys Gly Cys Val Gly Gly
 165 170 175
 Ala Lys Ser Ala Thr Ile Val Ala Asp Leu Ser Glu Thr Thr Ile Gln
 180 185 190
 Thr Pro Asp Gly Glu Glu Arg Leu Gly Gly Glu Ala His Ser Met Val
 195 200 205
 Trp Asp Pro Ser Gly Glu Arg Leu Ala Val Leu Met Lys Gly Lys Pro
 210 215 220
 Arg Val Gln Asp Gly Lys Pro Val Ile Leu Leu Phe Arg Thr Arg Asn
 225 230 235 240
 Ser Pro Val Phe Glu Leu Leu Pro Cys Gly Ile Ile Gln Gly Glu Pro
 245 250 255
 Gly Ala Gln Pro Gln Leu Ile Thr Phe His Leu Pro Ser Thr Lys Gly
 260 265 270
 Pro Cys Ser Val Trp Ala Gly Pro Gln Ala Glu Leu Pro Thr Ser Arg
 275 280 285
 Cys Thr Leu Ser Met Pro Ser Phe His Val Leu Ala Gln Cys Leu Gly
 290 295 300
 Gly Pro Arg Asn Pro Leu Leu Gly Val Glu Ala Leu Phe Met Thr Cys
 305 310 315 320
 Pro Ser Leu Leu Arg His Pro Gln Pro Leu Pro Leu Gly Thr Leu Ser
 325 330 335
 Gln Gly His His Leu Phe Cys Pro Thr Pro His Ile Pro Thr Ser Lys
 340 345 350
 Asn Lys Xaa
 355

<210> 169

<211> 90

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (90)

<223> Xaa equals stop translation

<400> 169

Met Cys Val Cys Tyr Phe Leu Val Phe Leu Gln Ile Trp Ala Arg Leu
 1 5 10 15

Ser His Leu Leu Val Trp Ile Tyr Pro Gly Ala Gly Leu Gln Pro Gly

99

20

25

30

Lys Gly His Pro Ala Gln Ser Leu Phe Pro His Glu His Cys His Leu
 35 40 45

Met Pro Gln His Ser Leu Thr Leu Lys Ile Leu Glu Glu Lys Leu Gly
 50 55 60

Gly Lys Gly Glu Ser Gly Ser Asn Phe Thr Phe Leu His Cys Lys Ile
 65 70 75 80

Leu Ala Thr Ser Ala Leu Asn Phe Ser Xaa
 85 90

<210> 170

<211> 59

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (59)

<223> Xaa equals stop translation

<400> 170

Met Val Leu Pro Phe Val Leu Leu Phe Arg Pro Asn Phe Ile Ser Val
 1 5 10 15

Leu His Pro Leu Phe Tyr Ser His Cys Leu Phe Leu Tyr Leu Ile Ser
 20 25 30

Pro Val His Ser Ser Ser Ile Ile Tyr Tyr Lys Pro Asp His Cys His
 35 40 45

Tyr Thr Pro Phe Ile Pro Gly Leu Leu Gln Xaa
 50 55

<210> 171

<211> 70

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (70)

<223> Xaa equals stop translation

<400> 171

Met Leu Leu Ser Lys Glu His Thr Ser Leu Gly Trp Leu Val Ile Phe
 1 5 10 15

Leu Thr Leu Ala Ser Gln Leu Ile Ser Tyr Gly Ser Arg Thr Gly Asn
 20 25 30

Ser Arg Cys Pro Pro Cys Leu Tyr Arg Thr Leu His Thr Val Ser Thr

35 40 100 45
 Ser His Val Leu Ser Ser Leu Phe Val Ser Thr Phe Ser Gly Asp Glu
 50 55 60
 Leu Val Trp Thr Thr Xaa
 65 70

<210> 172
 <211> 79
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (79)
 <223> Xaa equals stop translation

<400> 172
 Met Val Leu Asp Phe Lys Arg Ala Gly Ser Phe Phe Leu Ser Phe Leu
 1 5 10 15
 Trp Thr Arg Glu Ala Phe Ala Phe Ile Phe Thr Leu Pro Leu Leu Leu
 20 25 30
 Ser Leu Cys Arg Gly Lys Met Lys Asn Ser Pro Arg Ser Asp Leu Ser
 35 40 45
 Arg Leu Lys Lys Asn Val Phe Asn Ala Phe Leu Pro Cys Leu Val Pro
 50 55 60
 Arg Phe Ile Ser Asn Arg Gly Cys Pro Val Tyr Arg Ser Cys Xaa
 65 70 75

<210> 173
 <211> 174
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (150)
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>
 <221> SITE
 <222> (152)
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>
 <221> SITE
 <222> (174)
 <223> Xaa equals stop translation

<400> 173

101

Met Gly Val Pro Thr Ala Pro Glu Ala Gly Ser Trp Arg Trp Gly Ser
 1 5 10 15

Leu Leu Phe Ala Leu Phe Leu Ala Ala Ser Leu Asp Ile Thr Ala Ala
 20 25 30

Ala Leu Ala Thr Gly Ala Cys Ile Val Glu Ser Ser Ala Ser Pro Ser
 35 40 45

Ser Cys Ser Trp Ser Thr Ser Lys Gly Arg Gln Pro Pro Thr Ala Val
 50 55 60

Pro Arg Ser Trp Cys Gly Trp Thr Ala Thr Phe Lys Gly Leu Lys Thr
 65 70 75 80

Pro Ala Leu Lys Pro His His Leu Pro Arg Gly Tyr Pro Arg Pro Lys
 85 90 95

Ser Gly Thr Pro Cys Pro Met Trp Pro Ser Gly Ser Leu Leu Ser Leu
 100 105 110

Gly Gly Ile Cys Phe Arg Ser Pro Ala Pro Pro Cys Leu Leu Gln Ala
 115 120 125

Pro Glu Thr Ser Ser Ser His Pro Trp Thr Leu Ser Leu Thr Leu Gln
 130 135 140

Thr Leu Arg Ser Ser Xaa Pro Xaa Gly Gly Gln Trp Ala Val Val Ala
 145 150 155 160

Gly Ser Gly Ala Gly Ala Phe Glu Pro Gly Leu Ala Leu Xaa
 165 170

<210> 174

<211> 64

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (64)

<223> Xaa equals stop translation

<400> 174

Met Phe Val Leu Trp Val Phe Lys Ile Thr Tyr Ile Tyr Ile Leu Phe
 1 5 10 15

Ala Lys Asn Lys Ser Leu Ala Ser Cys Gln Met Ile Ala Lys Val Asp
 20 25 30

Leu Thr Phe Phe Val Ile Met Tyr Ile Phe Ile His Thr Pro Asn Thr
 35 40 45

Leu Ser Asp Phe Cys Tyr Phe Leu Gly Ser Thr Ala Leu Arg Leu Xaa
 50 55 60

<210> 175
 <211> 43
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (43)
 <223> Xaa equals stop translation

<400> 175
 Met Ile Ser Ala Gln Ser Ser Ile Ser Trp Ala Leu Ile Phe Ile Met
 1 5 10 15
 Ala Pro Ala Leu His Leu Val Leu Arg Phe Pro Ser Lys Phe Lys Pro
 20 25 30
 Glu Arg Lys Gly Glu Ala Arg Ser Pro Lys Xaa
 35 40

<210> 176
 <211> 114
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (114)
 <223> Xaa equals stop translation

<400> 176
 Met Trp Ile Ala Gly Pro Ser Trp Val Pro Leu Arg Tyr Val Val Trp
 1 5 10 15
 Leu Met His Leu Glu Arg Ile Cys Ala Leu His Asn Cys Arg Gly Asn
 20 25 30
 Met Leu Ser Trp Pro Leu Gln Ile Arg Val Ala Val Leu Gly Cys Cys
 35 40 45
 Thr Lys Thr Pro Ala Val Gly Phe Leu Gln Val Ala Gly Ser Pro His
 50 55 60
 Ser Cys Gln Asp Pro Gly Pro Cys Ser His Ser Ala Ala Ile Phe Pro
 65 70 75 80
 Pro Cys Glu Arg Gly Leu Cys Gly Asp Gly Pro Arg Cys Val Arg Gly
 85 90 95
 Cys Val His Cys His Arg Ser Leu Leu His Glu Pro Ala Trp Thr Gln
 100 105 110

SUBSTITUTE SHEET (RULE 26)

Gly Xaa

<210> 177

<211> 156

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (156)

<223> Xaa equals stop translation

<400> 177

Met	Ala	Ser	Ser	Leu	Ala	Phe	Leu	Leu	Leu	Asn	Phe	His	Val	Ser	Leu
1				5					10					15	

Leu	Leu	Val	Gln	Leu	Leu	Thr	Pro	Cys	Ser	Ala	Gln	Phe	Ser	Val	Leu
		20						25					30		

Gly	Pro	Ser	Gly	Pro	Ile	Leu	Ala	Met	Val	Gly	Glu	Asp	Ala	Asp	Leu
	35						40					45			

Pro	Cys	His	Leu	Phe	Pro	Thr	Met	Ser	Ala	Glu	Thr	Met	Glu	Leu	Lys
	50					55					60				

Trp	Val	Ser	Ser	Ser	Leu	Arg	Gln	Val	Val	Asn	Val	Tyr	Ala	Asp	Gly
65					70					75					80

Lys	Glu	Val	Glu	Asp	Arg	Gln	Ser	Ala	Pro	Tyr	Arg	Gly	Arg	Thr	Ser
				85					90					95	

Ile	Leu	Arg	Asp	Gly	Ile	Thr	Ala	Gly	Lys	Ala	Ala	Leu	Arg	Ile	His
		100						105						110	

Asn	Val	Thr	Ala	Ser	Asp	Ser	Gly	Lys	Tyr	Leu	Cys	Tyr	Phe	Gln	Asp
		115					120					125			

Gly	Asp	Phe	Tyr	Glu	Lys	Ala	Leu	Val	Glu	Leu	Lys	Val	Ala	Ala	Leu
	130					135					140				

Gly	Ser	Asn	Leu	His	Val	Gly	Ser	Glu	Gly	Leu	Xaa
145					150					155	

<210> 178

<211> 89

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (89)

<223> Xaa equals stop translation

<400> 178

104

Met Trp Pro Ser Gln Val Pro Leu Leu Ala Phe Cys Phe Leu Leu Val
 1 5 10 15

Lys Ser Thr Ser Asn Ile Asn Leu Pro Thr Pro Pro Ser Ser Leu
 20 25 30

Glu Asn Ser Ser Phe Val Val Ser Gln Arg Gly Asn Leu Ile Val Phe
 35 40 45

Gly Gly Gln Lys Lys Ala Thr Phe Arg Tyr His Phe Tyr Leu Asp Arg
 50 55 60

Met Pro Phe Tyr Ser Gln Ile Ser Val Tyr Phe Val Asn Gly Phe Arg
 65 70 75 80

Val Asn Gly Tyr Leu Cys Asn Asn Xaa
 85

<210> 179

<211> 197

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (197)

<223> Xaa equals stop translation

<400> 179

Met Ala Phe Arg Tyr Leu Ser Trp Ile Leu Phe Pro Leu Leu Gly Cys
 1 5 10 15

Tyr Ala Val Tyr Ser Leu Leu Tyr Leu Glu His Lys Gly Trp Tyr Ser
 20 25 30

Trp Val Leu Ser Met Leu Tyr Gly Phe Leu Leu Thr Phe Gly Phe Ile
 35 40 45

Thr Met Thr Pro Gln Leu Phe Ile Asn Tyr Lys Leu Lys Ser Val Ala
 50 55 60

His Leu Pro Trp Arg Met Leu Thr Tyr Lys Ala Leu Asn Thr Phe Ile
 65 70 75 80

Asp Asp Leu Phe Ala Phe Val Ile Lys Met Pro Val Met Tyr Arg Ile
 85 90 95

Gly Cys Leu Arg Asp Asp Val Val Phe Phe Ile Tyr Leu Tyr Gln Arg
 100 105 110

Trp Ile Tyr Arg Val Asp Pro Thr Arg Val Asn Glu Phe Gly Met Ser
 115 120 125

Gly Glu Asp Pro Thr Ala Ala Ala Pro Val Ala Glu Val Pro Thr Ala
 130 135 140

105

Ala Gly Ala Leu Thr Pro Thr Pro Ala Pro Thr Thr Thr Thr Ala Thr
 145 150 155 160

Arg Glu Glu Ala Ser Thr Ser Leu Pro Thr Lys Pro Thr Gln Gly Ala
 165 170 175

Ser Ser Ala Ser Glu Pro Gln Glu Ala Pro Pro Lys Pro Ala Glu Asp
 180 185 190

Lys Lys Lys Asp Xaa
 195

<210> 180

<211> 129

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (129)

<223> Xaa equals stop translation

<400> 180

Met Tyr Glu Cys Phe Leu Ser Leu Ser Leu Leu Lys Ser Cys Lys Ala
 1 5 10 15

Val Ser Gly Leu Met Cys Leu Leu Leu Pro Arg Leu Gly Leu Leu Leu
 20 25 30

Leu Leu Pro Ser Glu Arg Cys Phe Cys Trp Ile Pro Val Tyr Ser Leu
 35 40 45

Ile Thr Cys Leu Ala Glu Cys Ser Val Val Leu Arg Asp Pro Gly Phe
 50 55 60

Ala Gly Ala Phe Gln Val His Arg Arg Gln Ala Cys Phe Ser Thr Leu
 65 70 75 80

Arg Trp Ser Cys Leu Leu Leu Trp Trp Val Ser Arg Val Ser Ala Gly
 85 90 95

Arg Pro Leu Ile Gly Ser Pro His Met Met Ala Pro Ser Thr Phe Cys
 100 105 110

Pro Thr Val Arg Gly Pro Gly Thr Cys Ala Ser Ser Asp Pro Asp Gly
 115 120 125

Xaa

<210> 181

<211> 155

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (155)

<223> Xaa equals stop translation

<400> 181

Met Pro Ala Glu Lys Arg Ile Phe Gly Ala Val Leu Leu Phe Ser Trp
 1 5 10 15

Thr Val Tyr Leu Trp Glu Thr Phe Leu Ala Gln Arg Gln Arg Arg Ile
 20 25 30

Tyr Lys Thr Thr Thr His Val Pro Pro Glu Leu Gly Gln Ile Met Asp
 35 40 45

Ser Glu Thr Phe Glu Lys Ser Arg Leu Tyr Gln Leu Asp Lys Ser Thr
 50 55 60

Phe Ser Phe Trp Ser Gly Leu Tyr Ser Glu Thr Glu Gly Thr Leu Asn
 65 70 75 80

Leu Leu Phe Gly Gly Ile Pro Tyr Leu Trp Arg Leu Ser Gly Arg Phe
 85 90 95

Cys Gly Tyr Ala Gly Phe Gly Pro Glu Tyr Glu Ile Thr Gln Ser Leu
 100 105 110

Val Phe Leu Leu Leu Ala Thr Leu Phe Ser Ala Leu Thr Gly Val Pro
 115 120 125

Trp Ser Leu Tyr Asn Thr Phe Val Ile Lys Lys Thr Trp Leu Gln Ser
 130 135 140

Thr Asp Phe Gly Val Leu His Met Glu Ile Xaa
 145 150 155

<210> 182

<211> 107

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (107)

<223> Xaa equals stop translation

<400> 182

Met Ser Leu Ser Trp Met Val His Leu Leu Gly Leu Pro Asn Gly Thr
 1 5 10 15

Val Trp Tyr Leu Pro Phe Val Cys Phe Thr Arg Gly Ser Pro Met Gly
 20 25 30

Gly Gly Ser Gly Gln Trp Arg Trp Asp Arg Lys Phe Ser Lys Thr Leu
 35 40 45

107

Leu Gly Asn Leu Phe Val Ala Phe Lys Glu Met Cys Gly Glu Asp Ile
 50 55 60

Trp Met Leu Ala Ala Ile Leu Glu Leu Arg Thr Gln Glu Trp Trp Lys
 65 70 75 80

Gly Arg Arg Asn Arg Val Phe Val Ala Val Val Lys Leu Leu Lys Phe
 85 90 95

Pro Ser Cys Gln Ala Ser Cys Tyr Met Arg Xaa
 100 105

<210> 183

<211> 48

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (48)

<223> Xaa equals stop translation

<400> 183

Met Ile Asn Glu Trp Cys Phe Lys Leu Leu Ser Leu Trp Ser Phe Ala
 1 5 10 15

Tyr Ser Asn Cys Lys Leu Ile His Lys Cys Lys Phe Val Phe Leu Lys
 20 25 30

Lys Lys Lys Thr Gly Lys Glu Val Ser Val Lys Gly Ser Lys Leu Xaa
 35 40 45

<210> 184

<211> 127

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (127)

<223> Xaa equals stop translation

<400> 184

Met Trp Leu Gly Ser Trp Leu Thr Ser Leu Leu Leu Ser Pro Tyr Gly
 1 5 10 15

Ser Gly Trp Glu Lys Val Pro Cys Cys Val Thr Gly His Leu Arg Ser
 20 25 30

Cys Ser Cys Cys Leu Leu Gly Leu Ala Gly Val Gln Ser Asp His Phe
 35 40 45

108

Ser Glu Gly Phe Phe Ser Glu Tyr Ser Ser Asp Val Leu Pro Trp Gly
 50 55 60

Arg Arg Ser Phe Leu Pro Gln Gly Asp Ala Ser Leu Leu Ala Cys Glu
 65 70 75 80

Cys Phe Leu His Leu Gln Val Val Trp Gly Gln Phe Cys Leu Leu Glu
 85 90 95

Ala Trp Ala Gly Phe Thr Glu Gly Ser Met Pro Ala Pro Ser Cys Arg
 100 105 110

Val His Phe Trp Cys Arg Val Asn Thr Cys Pro Phe Met Ser Xaa
 115 120 125

<210> 185
 <211> 87
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (87)
 <223> Xaa equals stop translation

<400> 185
 Met Leu Cys Gly Tyr Val Ile Asn Asn Ile Trp Leu Ile Phe Thr Tyr
 1 5 10 15

Phe Ile Cys Ile Tyr Ile Ser Arg Ser Tyr Ile Tyr Ile Thr Gln Glu
 20 25 30

Thr Gln Val Ile Tyr Ile Cys Gln Glu Met Tyr Asp Tyr Phe Gly Glu
 35 40 45

Asn Gly Pro Lys Cys Glu Lys Asp Ile Lys Lys Thr Lys Lys Thr Lys
 50 55 60

Lys Lys His Tyr Phe Pro Leu Arg Asn Ile Leu Tyr Ile Ser Lys Glu
 65 70 75 80

Glu Lys Leu Lys Asp Ile Xaa
 85

<210> 186
 <211> 58
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (58)
 <223> Xaa equals stop translation

<400> 186

109

Met Ile Val Ser Tyr Arg Ile Val Ser Leu Pro Ser Ser Val Leu Cys
 1 5 10 15

Leu Phe Ile Pro Pro Phe Leu Leu Ile Phe Tyr Cys Leu His Ser Phe
 20 25 30

Val Phe Ser Gln Met Leu Tyr Ser Trp Asn Tyr His Val Thr Phe Gln
 35 40 45

Met Ala Phe Ser Leu Ile Ile Cys Val Xaa
 50 55

<210> 187

<211> 69

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (69)

<223> Xaa equals stop translation

<400> 187

Met Val Ala Ser Gln Ala Trp Trp Leu Ser Asn Leu Trp His Leu Trp
 1 5 10 15

Glu Val Gly Ser Ala Gln Gly Leu Pro Leu Asp Pro Pro Ala Leu Ala
 20 25 30

Pro Tyr Leu Pro Trp Ala Leu Arg Trp Pro Cys Phe Ser Gly Phe Ala
 35 40 45

Ser Leu Ala Gly Ala Leu Val Leu Ala His Ser Leu Pro Thr Ala Trp
 50 55 60

Pro Gly Ser Ser Xaa
 65

<210> 188

<211> 48

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (48)

<223> Xaa equals stop translation

<400> 188

Met Tyr Leu Phe Leu Leu Cys Cys Phe Ile Ser Glu His Cys Ala Gln
 1 5 10 15

His Ser Phe Pro His Thr Cys Pro Asn Trp Lys Thr Arg Val Leu Ser
 20 25 30

110

Phe Pro Leu His Pro Cys Pro His Leu Ile His Pro Asn Asn Thr Xaa
 35 40 45

<210> 189

<211> 51

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (5)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (51)

<223> Xaa equals stop translation

<400> 189

Met Leu Ser Ser Xaa Tyr Val Pro Met Cys Gln His Phe Ile Tyr Pro
 1 5 10 15

Val Leu Trp Val Leu Val His Phe Phe Ser Phe Ile Gln Ile Gln Lys
 20 25 30

Asn Thr Asp Gly Ser Asn Val Lys Leu Thr Arg Asn Pro Gly Thr Phe
 35 40 45

Ile Ser Xaa
 50

<210> 190

<211> 56

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (56)

<223> Xaa equals stop translation

<400> 190

Met Ala Val Arg Val Leu Trp Gly Gly Leu Ser Leu Leu Arg Val Leu
 1 5 10 15

Trp Cys Leu Leu Pro Gln Thr Gly Tyr Val His Pro Asp Glu Phe Phe
 20 25 30

Gln Ser Pro Glu Val Met Ala Gly Lys Thr Pro His Val Trp Leu Arg
 35 40 45

Gln Ala Ala Ala Glu Ser Ala Xaa

50

55

<210> 191
 <211> 127
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (127)
 <223> Xaa equals stop translation

<400> 191
 Met Cys Ser Ser Phe Pro Arg Met Ala Leu Cys Ala Leu Trp Met Trp
 1 5 10 15
 Pro Ser Val Lys Ser Ser Val Pro Leu Pro Leu Arg Glu Pro Phe Leu
 20 25 30
 Trp Arg Ser Pro Gly Ser Gln Cys Leu Leu Cys Leu Gln Thr Ile His
 35 40 45
 Val Ser Cys Ser Glu Ala Cys Pro Leu Leu Glu Asn Ile Ser Lys Asn
 50 55 60
 Cys Thr Ile Pro Gln Arg Asp Leu Asp Asn Met Ala Phe Pro Gln Ala
 65 70 75 80
 Leu Pro Leu Glu Lys Arg Cys Glu Arg Phe Leu Gln Lys Ser Tyr Arg
 85 90 95
 Lys Leu Glu Lys Asn Pro Glu Lys Glu Glu Glu His Trp Ala Arg Leu
 100 105 110
 Gln Arg Tyr Ser Leu Ser Leu Gln Arg Glu Asn Phe Lys Lys Xaa
 115 120 125

<210> 192
 <211> 70
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (70)
 <223> Xaa equals stop translation

<400> 192
 Met Pro Phe Gln Leu Pro Leu Gln Leu Leu Leu Arg Leu Ile Cys
 1 5 10 15
 Glu Phe Phe Leu Ala Pro Ala Leu Asn Cys Asn Leu Thr Gly Thr Val
 20 25 30
 Ile Phe Phe Thr Leu Met Ile Ser Leu Gln Leu Met Ile Phe Phe Thr

112

35

40

45

Leu Gln Phe Ala Asp Gly Phe Gln Ile Gly Val Asp Leu Gln Leu Ser
 50 55 60

Glu Leu Asn Ile Leu Xaa
 65 70

<210> 193
 <211> 71
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (71)
 <223> Xaa equals stop translation

<400> 193
 Met Ile Ser Gly Val Leu Ile Phe Asn Leu Ile Ala Ser Ser Trp Val
 1 5 10 15

Leu Cys Phe Pro Leu Cys Asp Leu Ser Cys Gln Lys Thr Leu Arg Ile
 20 25 30

Phe Phe Ala Ser Phe Phe His Ala Val Cys Val His Val Ser Cys Thr
 35 40 45

Ser Trp Gln Pro Leu Val Leu Phe Ile Lys Trp Trp Val Val Gly Cys
 50 55 60

Ser Pro Ala Val Ser Leu Xaa
 65 70

<210> 194
 <211> 130
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (130)
 <223> Xaa equals stop translation

<400> 194
 Met His Val Leu Pro Leu Leu Leu Ser Leu Leu Leu Leu Leu Leu
 1 5 10 15

Leu Ser Ala Ser Phe Val Thr Phe Ser Thr Pro Thr Ser Ser Arg Asn
 20 25 30

Ser Ser Cys Pro Asp Cys Glu Ser Leu Asn Thr Gly Leu Pro Ser Leu
 35 40 45

Met Met Phe Gly Gly Ser Leu Leu Lys Trp Val Gln Asn Thr His Gly

50 55 113 60
 Val Glu Ser Leu Leu Ser Ser Ala Lys Val Arg Leu Leu Pro Pro Ala
 65 70 75 80
 Leu Gly Val Leu Phe Pro Arg Leu His Pro Gly Thr Leu Thr Leu Val
 85 90 95
 Phe Leu Leu Ile Pro Phe Leu Thr Val Ser Ser Ser Thr Ser Asp Val
 100 105 110
 Leu Ser Ser Leu Glu Ser Pro Lys Leu Ser Val Thr Ile Phe His Tyr
 115 120 125
 Cys Xaa
 130

<210> 195
 <211> 55
 <212> PRT
 <213> Homo sapiens
 <220>
 <221> SITE
 <222> (55)
 <223> Xaa equals stop translation

<400> 195
 Met Pro Trp Ile Leu Met Leu Leu Phe Thr Met Gly Gln Gly Val Val
 1 5 10 15
 Ile Leu Ala Phe Arg Ser Cys Leu Glu Ala Glu Val Arg Gly Val Pro
 20 25 30
 Gly Arg Gly Asn Arg Ser Gly Val Lys Thr Val Val Glu Ala Pro Ala
 35 40 45
 Val Phe Ala Lys Arg Pro Xaa
 50 55

<210> 196
 <211> 80
 <212> PRT
 <213> Homo sapiens
 <220>
 <221> SITE
 <222> (80)
 <223> Xaa equals stop translation

<400> 196
 Met Ala Ala Phe Phe Ala Leu Ala Ala Leu Val Gln Val Val Tyr Thr
 1 5 10 15
 Ile Pro Ala Val Leu Thr Leu Leu Val Gly Leu Asn Pro Glu Val Thr

114

20

25

30

Gly Asn Val Ile Trp Lys Ser Ile Ser Ala Ile His Ile Leu Phe Cys
 35 40 45

Thr Val Trp Ala Val Gly Leu Ala Ser Tyr Leu Leu His Arg Thr Gln
 50 55 60

Gln Asn Ile Leu His Glu Glu Gly Arg Ser Cys Leu Val Trp Xaa
 65 70 75 80

<210> 197
 <211> 42
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (42)
 <223> Xaa equals stop translation

<400> 197
 Met Lys His Met Asn Thr Leu Pro Ile Phe Ser Ser Leu Ile Ser Phe
 1 5 10 15

Leu Pro Ala Val Ser Ala Gly Arg Ser Ala Ile Thr Thr Leu Cys Asn
 20 25 30

Ile Thr Glu Gln Leu Glu Val Leu Gly Xaa
 35 40

<210> 198
 <211> 197
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (197)
 <223> Xaa equals stop translation

<400> 198
 Met Lys Tyr Leu Arg His Arg Arg Pro Asn Ala Thr Leu Ile Leu Ala
 1 5 10 15

Ile Gly Ala Phe Thr Leu Leu Leu Phe Ser Leu Leu Val Ser Pro Pro
 20 25 30

Thr Cys Lys Val Gln Glu Gln Pro Pro Ala Ile Pro Glu Ala Leu Ala
 35 40 45

Trp Pro Thr Pro Pro Thr Arg Pro Ala Pro Ala Pro Cys His Ala Asn

50 55 115 60
 Thr Ser Met Val Thr His Pro Asp Phe Ala Thr Gln Pro Gln His Val
 65 70 75 80
 Gln Asn Phe Leu Leu Tyr Arg His Cys Arg His Phe Pro Leu Leu Gln
 85 90 95
 Asp Val Pro Pro Ser Lys Cys Ala Gln Pro Val Phe Leu Leu Val
 100 105 110
 Ile Lys Ser Ser Pro Ser Asn Tyr Val Arg Arg Glu Leu Leu Arg Arg
 115 120 125
 Thr Trp Gly Arg Glu Arg Lys Val Arg Gly Leu Gln Leu Arg Leu Leu
 130 135 140
 Phe Leu Val Gly Thr Ala Ser Asn Pro His Glu Ala Arg Lys Val Asn
 145 150 155 160
 Arg Leu Leu Glu Leu Glu Ala Gln Thr His Gly Asp Ile Leu Gln Trp
 165 170 175
 Asp Phe His Asp Ser Phe Phe Asn Leu Thr Leu Lys Gln Val Arg Trp
 180 185 190
 Thr Gly Val Thr Xaa
 195

 <210> 199
 <211> 124
 <212> PRT
 <213> Homo sapiens

 <220>
 <221> SITE
 <222> (124)
 <223> Xaa equals stop translation

 <400> 199
 Met Lys Leu Leu Leu Ala Leu Pro Met Leu Val Leu Leu Pro Gln
 1 5 10 15
 Val Ile Pro Ala Tyr Ser Gly Glu Lys Lys Cys Trp Asn Arg Ser Gly
 20 25 30
 His Cys Arg Lys Gln Cys Lys Asp Gly Glu Ala Val Lys Asp Thr Cys
 35 40 45
 Lys Asn Leu Arg Ala Cys Cys Ile Pro Ser Asn Glu Asp His Arg Arg
 50 55 60
 Val Pro Ala Thr Ser Pro Thr Pro Leu Ser Asp Ser Thr Pro Gly Ile
 65 70 75 80
 Ile Asp Asp Ile Leu Thr Val Arg Phe Thr Thr Asp Tyr Phe Glu Val

85 116 95
 Ser Ser Lys Lys Asp Met Val Glu Glu Ser Glu Ala Gly Arg Gly Thr
 100 105 110
 Glu Thr Ser Leu Pro Asn Val His His Ser Ser Xaa
 115 120

<210> 200
 <211> 549
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (132)
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>
 <221> SITE
 <222> (398)
 <223> Xaa equals any of the naturally occurring L-amino acids

<400> 200
 Met Gly Asn Ala Cys Ile Pro Leu Lys Arg Ile Ala Tyr Phe Leu Cys
 1 5 10 15
 Leu Leu Ser Ala Leu Leu Leu Thr Glu Gly Lys Lys Pro Ala Lys Pro
 20 25 30
 Lys Cys Pro Ala Val Cys Thr Cys Thr Lys Asp Asn Ala Leu Cys Glu
 35 40 45
 Asn Ala Arg Ser Ile Pro Arg Thr Val Pro Pro Asp Val Ile Ser Leu
 50 55 60
 Ser Phe Val Arg Ser Gly Phe Thr Glu Ile Ser Glu Gly Ser Phe Leu
 65 70 75 80
 Phe Thr Pro Ser Leu Gln Leu Leu Leu Phe Thr Ser Asn Ser Phe Asp
 85 90 95
 Val Ile Ser Asp Asp Ala Phe Ile Gly Leu Pro His Leu Glu Tyr Leu
 100 105 110
 Phe Ile Glu Asn Asn Asn Ile Lys Ser Ile Ser Arg His Thr Phe Arg
 115 120 125
 Gly Leu Lys Xaa Leu Ile His Leu Ser Leu Ala Asn Asn Asn Leu Gln
 130 135 140
 Thr Leu Pro Lys Asp Ile Phe Lys Gly Leu Asp Ser Leu Thr Asn Val
 145 150 155 160
 Asp Leu Arg Gly Asn Ser Phe Asn Cys Asp Cys Lys Leu Lys Trp Leu
 165 170 175

Val Glu Trp Leu Gly His Thr Asn Ala Thr Val Glu Asp Ile Tyr Cys
 180 185 190
 Glu Gly Pro Pro Glu Tyr Lys Lys Arg Lys Ile Asn Ser Leu Ser Ser
 195 200 205
 Lys Asp Phe Asp Cys Ile Ile Thr Glu Phe Ala Lys Ser Gln Asp Leu
 210 215 220
 Pro Tyr Gln Ser Leu Ser Ile Asp Thr Phe Ser Tyr Leu Asn Asp Glu
 225 230 235 240
 Tyr Val Val Ile Ala Gln Pro Phe Thr Gly Lys Cys Ile Phe Leu Glu
 245 250 255
 Trp Asp His Val Glu Lys Thr Phe Arg Asn Tyr Asp Asn Ile Thr Gly
 260 265 270
 Thr Ser Thr Val Val Cys Lys Pro Ile Val Ile Glu Thr Gln Leu Tyr
 275 280 285
 Val Ile Val Ala Gln Leu Phe Gly Gly Ser His Ile Tyr Lys Arg Asp
 290 295 300
 Ser Phe Ala Asn Lys Phe Ile Lys Ile Gln Asp Ile Glu Ile Leu Lys
 305 310 315 320
 Ile Arg Lys Pro Asn Asp Ile Glu Thr Phe Lys Ile Glu Asn Asn Trp
 325 330 335
 Tyr Phe Val Val Ala Asp Ser Ser Lys Ala Gly Phe Thr Thr Ile Tyr
 340 345 350
 Lys Trp Asn Gly Asn Gly Phe Tyr Ser His Gln Ser Leu His Ala Trp
 355 360 365
 Tyr Arg Asp Thr Asp Val Glu Tyr Leu Glu Ile Val Arg Thr Pro Gln
 370 375 380
 Thr Leu Arg Thr Pro His Leu Ile Leu Ser Ser Ser Ser Xaa Arg Pro
 385 390 395 400
 Val Ile Tyr Gln Trp Asn Lys Ala Thr Gln Leu Phe Thr Asn Gln Thr
 405 410 415
 Asp Ile Pro Asn Met Glu Asp Val Tyr Ala Val Lys His Phe Ser Val
 420 425 430
 Lys Gly Asp Val Tyr Ile Cys Leu Thr Arg Phe Ile Gly Asp Ser Lys
 435 440 445
 Val Met Lys Trp Gly Gly Ser Ser Phe Gln Asp Ile Gln Arg Met Pro
 450 455 460
 Ser Arg Gly Ser Met Val Phe Gln Pro Leu Gln Ile Asn Asn Tyr Gln
 465 470 475 480

Tyr Ala Ile Leu Gly Ser Asp Tyr Ser Phe Thr Gln Val Tyr Asn Trp
 485 490 495

Asp Ala Glu Lys Ala Lys Phe Val Lys Phe Gln Glu Leu Asn Val Gln
 500 505 510

Ala Pro Arg Ser Phe Thr His Val Ser Ile Asn Lys Arg Asn Phe Leu
 515 520 525

Phe Ala Ser Ser Phe Lys Gly Asn Thr Gln Ile Tyr Lys His Val Ile
 530 535 540

Val Asp Leu Ser Ala
 545

<210> 201

<211> 488

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (344)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (416)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (429)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (430)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 201

Met Ile Leu Ser Leu Leu Phe Ser Leu Gly Gly Pro Leu Gly Trp Gly
 1 5 10 15

Leu Leu Gly Ala Trp Ala Gln Ala Ser Ser Thr Ser Leu Ser Asp Leu
 20 25 30

Gln Ser Ser Arg Thr Pro Gly Val Trp Lys Ala Glu Ala Glu Asp Thr
 35 40 45

Ser Lys Asp Pro Val Gly Arg Asn Trp Cys Pro Tyr Pro Met Ser Lys
 50 55 60

Leu Val Thr Leu Leu Ala Leu Cys Lys Thr Glu Lys Phe Leu Ile His
 65 70 75 80

Ser Gln Gln Pro Cys Pro Gln Gly Ala Pro Asp Cys Gln Lys Val Lys
 85 90 95
 Val Met Tyr Arg Met Ala His Lys Pro Val Tyr Gln Val Lys Gln Lys
 100 105 110
 Val Leu Thr Ser Leu Ala Trp Arg Cys Cys Pro Gly Tyr Thr Gly Pro
 115 120 125
 Asn Cys Glu His His Asp Ser Met Ala Ile Pro Glu Pro Ala Asp Pro
 130 135 140
 Gly Asp Ser His Gln Glu Pro Gln Asp Gly Pro Val Ser Phe Lys Pro
 145 150 155 160
 Gly His Leu Ala Ala Val Ile Asn Glu Val Glu Val Gln Gln Glu Gln
 165 170 175
 Gln Glu His Leu Leu Gly Asp Leu Gln Asn Asp Val His Arg Val Ala
 180 185 190
 Asp Ser Leu Pro Gly Leu Trp Lys Ala Leu Pro Gly Asn Leu Thr Ala
 195 200 205
 Ala Val Met Glu Ala Asn Gln Thr Gly His Glu Phe Pro Asp Arg Ser
 210 215 220
 Leu Glu Gln Val Leu Leu Pro His Val Asp Thr Phe Leu Gln Val His
 225 230 235 240
 Phe Ser Pro Ile Trp Arg Ser Phe Asn Gln Ser Leu His Ser Leu Thr
 245 250 255
 Gln Ala Ile Arg Asn Leu Ser Leu Asp Val Glu Ala Asn Arg Gln Ala
 260 265 270
 Ile Ser Arg Val Gln Asp Ser Ala Val Ala Arg Ala Asp Phe Gln Glu
 275 280 285
 Leu Gly Ala Lys Phe Glu Ala Lys Val Gln Glu Asn Thr Gln Arg Val
 290 295 300
 Gly Gln Leu Arg Gln Asp Val Glu Glu Arg Leu His Ala Gln His Phe
 305 310 315 320
 Thr Leu His Arg Ser Ile Ser Glu Leu Gln Ala Asp Val Asp Thr Lys
 325 330 335
 Leu Lys Arg Leu His Lys Ala Xaa Glu Ala Pro Gly Thr Asn Gly Ser
 340 345 350
 Leu Val Leu Ala Thr Pro Gly Ala Gly Ala Arg Pro Glu Pro Asp Ser
 355 360 365
 Leu Gln Ala Arg Leu Gly Gln Leu Gln Arg Asn Leu Ser Glu Leu His
 370 375 380

Met Thr Thr Ala Arg Arg Glu Glu Glu Leu Gln Tyr Thr Leu Glu Asp
 385 390 395 400

Met Arg Ala Thr Leu Thr Arg His Val Asp Glu Ile Lys Glu Leu Xaa
 405 410 415

Ser Glu Ser Asp Glu Thr Phe Asp Gln Ile Ser Lys Xaa Xaa Arg Gln
 420 425 430

Val Glu Glu Leu Gln Val Asn His Thr Ala Leu Arg Glu Leu Arg Val
 435 440 445

Ile Leu Met Glu Lys Ser Leu Ile Met Glu Glu Asn Lys Glu Glu Val
 450 455 460

Glu Arg Gln Leu Leu Glu Leu Asn Leu Thr Leu Gln His Leu Gln Gly
 465 470 475 480

Gly Met Pro Thr Ser Ser Thr
 485

<210> 202

<211> 86

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (86)

<223> Xaa equals stop translation

<400> 202

Met Ala His Gly Pro Gln Ser Leu Trp Ser Leu Gly Phe Thr Val Thr
 1 5 10 15

Leu Thr Phe Glu Leu Pro Val Gly Cys Val Leu Gly Arg Ile Cys His
 20 25 30

Pro Ile Gln Ala Cys Asn Thr Gly Leu Met Thr Pro Thr Pro Gln Gly
 35 40 45

Pro Cys Arg Thr Glu Met Met Ser Asn Asp Lys Pro Trp Leu Pro Ala
 50 55 60

Asn Ala Pro Ala His Ile Ser Leu Pro Gly Ala Arg Leu Thr Ser Thr
 65 70 75 80

Cys Ala Pro Gly Leu Xaa
 85

<210> 203

<211> 400

<212> PRT

<213> Homo sapiens

<220>
 <221> SITE
 <222> (400)
 <223> Xaa equals stop translation

<400> 203

```

Met Ala Ile His Lys Ala Leu Val Met Cys Leu Gly Leu Pro Leu Phe
  1              5              10              15

Leu Phe Pro Gly Ala Trp Ala Gln Gly His Val Pro Pro Gly Cys Ser
          20              25              30

Gln Gly Leu Asn Pro Leu Tyr Tyr Asn Leu Cys Asp Arg Ser Gly Ala
      35              40              45

Trp Gly Ile Val Leu Glu Ala Val Ala Gly Ala Gly Ile Val Thr Thr
      50              55              60

Phe Val Leu Thr Ile Ile Leu Val Ala Ser Leu Pro Phe Val Gln Asp
      65              70              75              80

Thr Lys Lys Arg Ser Leu Leu Gly Thr Gln Val Phe Phe Leu Leu Gly
          85              90              95

Thr Leu Gly Leu Phe Cys Leu Val Phe Ala Cys Val Val Lys Pro Asp
      100              105              110

Phe Ser Thr Cys Ala Ser Arg Arg Phe Leu Phe Gly Val Leu Phe Ala
      115              120              125

Ile Cys Phe Ser Cys Leu Ala Ala His Val Phe Ala Leu Asn Phe Leu
      130              135              140

Ala Arg Lys Asn His Gly Pro Arg Gly Trp Val Ile Phe Thr Val Ala
      145              150              155              160

Leu Leu Leu Thr Leu Val Glu Val Ile Ile Asn Thr Glu Trp Leu Ile
      165              170              175

Ile Thr Leu Val Arg Gly Ser Gly Glu Gly Gly Pro Gln Gly Asn Ser
      180              185              190

Ser Ala Gly Trp Ala Val Ala Ser Pro Cys Ala Ile Ala Asn Met Asp
      195              200              205

Phe Val Met Ala Leu Ile Tyr Val Met Leu Leu Leu Leu Gly Ala Phe
      210              215              220

Leu Gly Ala Trp Pro Ala Leu Cys Gly Arg Tyr Lys Arg Trp Arg Lys
      225              230              235              240

His Gly Val Phe Val Leu Leu Thr Thr Ala Thr Ser Val Ala Ile Trp
          245              250              255

Val Val Trp Ile Val Met Tyr Thr Tyr Gly Asn Lys Gln His Asn Ser
      260              265              270
  
```

SUBSTITUTE SHEET (RULE 26)

Pro Thr Trp Asp Asp Pro Thr Leu Ala Ile Ala Leu Ala Ala Asn Ala
 275 280 285

Trp Ala Phe Val Leu Phe Tyr Val Ile Pro Glu Val Ser Gln Val Thr
 290 295 300

Lys Ser Ser Pro Glu Gln Ser Tyr Gln Gly Asp Met Tyr Pro Thr Arg
 305 310 315 320

Gly Val Gly Tyr Glu Thr Ile Leu Lys Glu Gln Lys Gly Gln Ser Met
 325 330 335

Phe Val Glu Asn Lys Ala Phe Ser Met Asp Glu Pro Val Ala Ala Lys
 340 345 350

Arg Pro Val Ser Pro Tyr Ser Gly Tyr Asn Gly Gln Leu Leu Thr Ser
 355 360 365

Val Tyr Gln Pro Thr Glu Met Ala Leu Met His Lys Val Pro Ser Glu
 370 375 380

Glu Leu Thr Thr Ser Ser Ser His Gly Pro Pro Pro Thr Ala Arg Xaa
 385 390 395 400

<210> 204
 <211> 195
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (195)
 <223> Xaa equals stop translation

<400> 204
 Met Ser Thr Ala Phe Cys Pro Ile His Ser Ser Leu Gly Ser Met Val
 1 5 10 15

Met Cys Leu Cys Ile Leu Ser Pro Leu Cys Ile Ala Ser Lys Ser Leu
 20 25 30

Arg Val Cys Thr Lys Ser Tyr Met Glu Gly His Gly Lys Thr Arg Val
 35 40 45

Pro Val Val Leu Val Gly Asn Lys Ala Asp Leu Ser Pro Glu Arg Glu
 50 55 60

Val Gln Ala Val Glu Gly Lys Lys Leu Ala Glu Ser Trp Gly Ala Thr
 65 70 75 80

Phe Met Glu Ser Ser Ala Arg Glu Asn Gln Leu Thr Gln Gly Ile Phe
 85 90 95

Thr Lys Val Ile Gln Glu Ile Ala Arg Val Gly Glu Phe Leu Trp Ala
 100 105 110
 Arg Ala Ser Leu Pro Ser His Val Ser Pro Trp Val Trp Gly Asn Cys
 115 120 125
 Leu Ala Ser Ala Pro Gly Thr Cys His Val Pro Val Gly Gly Arg Ser
 130 135 140
 Ser Gly Leu His Gly Tyr Gly Cys Gln Leu Cys Ser Trp Pro Leu Asp
 145 150 155 160
 Thr Gln Cys Gly Ile Leu Met Phe Ala His Phe Pro Gln Ala Pro Val
 165 170 175
 Ala Trp Met Ser Met Phe Thr Lys Gly Gln Gly Pro Leu Met Asp Thr
 180 185 190
 Gly Leu Xaa
 195

<210> 205
 <211> 57
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (57)
 <223> Xaa equals stop translation

<400> 205
 Met Pro Leu Glu Glu Ser Phe Glu Ile Val Leu Lys Leu Val Pro Leu
 1 5 10 15
 Leu Gly Leu Glu Leu Phe Phe Phe Leu Phe Ile Ile Asn Gly Tyr Ile
 20 25 30
 Asn Val Tyr Cys Pro Ser Gln Tyr Phe Ile Tyr Ala Lys Asp Ser Leu
 35 40 45
 Ala Gly Leu Ala Leu Ile Pro Gln Xaa
 50 55

<210> 206
 <211> 73
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (73)
 <223> Xaa equals stop translation

124

<400> 206

Met Ile Val Ile Tyr Leu Thr Leu Thr Trp Thr Phe Leu Ile Asn Leu
 1 5 10 15

Leu Ala Cys Pro Leu Tyr His Leu Pro Gln Met Gln Lys Lys Ala Lys
 20 25 30

Pro Glu Thr Lys Lys Ala Lys Pro Glu Thr Lys Glu Thr Ile Gln Arg
 35 40 45

Gln Arg Asn Leu Phe Leu Val Leu Leu Lys Gln Leu Ala Gly Lys Lys
 50 55 60

Cys Ser Ala Leu Phe Leu Ile Val Xaa
 65 70

<210> 207

<211> 85

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (85)

<223> Xaa equals stop translation

<400> 207

Met Val Trp Cys Gln Cys Leu Cys Pro Leu Cys Ala Cys Trp Glu Glu
 1 5 10 15

Ala Gln Ala Leu Trp Trp Pro Pro Leu Cys Thr Trp Pro Gly Glu Ala
 20 25 30

Arg Gly Ser Gly Ala Ser Leu Arg Leu Arg Pro Pro Leu Gln Asn Lys
 35 40 45

Leu Ser Pro Gly Val Cys Leu Ser Leu Phe Leu Ser Pro Glu Arg Asn
 50 55 60

Ala Gly Val Pro Glu Ala Ser Leu Gln Thr Lys His Pro Cys Thr Ser
 65 70 75 80

Tyr Gly Ser Gly Xaa
 85

<210> 208

<211> 195

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (195)

<223> Xaa equals stop translation

125

<400> 208

Met Trp Val Ser Leu Tyr Phe Gly Ile Leu Gly Leu Cys Ser Val Ile
 1 5 10 15

Thr Gly Gly Cys Ile Ile Phe Leu His Trp Arg Lys Asn Leu Arg Arg
 20 25 30

Glu Glu His Ala Gln Gln Trp Val Glu Val Met Arg Ala Ala Thr Phe
 35 40 45

Thr Tyr Ser Pro Leu Leu Tyr Trp Ile Asn Lys Arg Arg Arg Tyr Gly
 50 55 60

Met Asn Ala Ala Ile Asn Thr Gly Pro Ala Pro Ala Val Thr Lys Thr
 65 70 75 80

Glu Thr Glu Val Gln Asn Pro Asp Val Leu Trp Asp Leu Asp Ile Pro
 85 90 95

Glu Gly Arg Ser His Ala Asp Gln Asp Ser Asn Pro Lys Ala Glu Ala
 100 105 110

Pro Ala Pro Leu Gln Pro Ala Leu Gln Leu Ala Pro Gln Gln Pro Gln
 115 120 125

Ala Arg Ser Pro Phe Pro Leu Pro Ile Phe Gln Glu Val Pro Phe Ala
 130 135 140

Pro Pro Leu Cys Asn Leu Pro Pro Leu Leu Asn His Ser Val Ser Tyr
 145 150 155 160

Pro Leu Ala Thr Cys Pro Glu Arg Asn Val Leu Phe His Ser Leu Leu
 165 170 175

Asn Leu Ala Gln Glu Asp His Ser Phe Asn Ala Lys Pro Phe Pro Ser
 180 185 190

Glu Leu Xaa
 195

<210> 209

<211> 42

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (42)

<223> Xaa equals stop translation

<400> 209

Met Leu Gln Arg Gly Gln His Leu Tyr Leu Val Val Phe Leu Met Val
 1 5 10 15

Ser Phe Ile Pro Leu Leu Asn Pro Lys Gln Asp Leu Lys Lys Leu Lys
 20 25 30

Lys Asn Arg Thr Val Arg Asn His Phe Xaa
 35 40

<210> 210
 <211> 282
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (282)
 <223> Xaa equals stop translation

<400> 210
 Met Ser Ile Leu Thr Met Ile Ser Ser Trp Pro Phe Ser Arg Val Val
 1 5 10 15
 Arg Phe Trp Phe Leu His Gln Met Val Leu Asp Leu Cys Leu Gly Gln
 20 25 30
 Gly Val Pro Gln Gln Asn Leu Gly Lys Pro Lys Gly Lys Lys Leu
 35 40 45
 Ser Ser Val Arg Gln Lys Phe Asp His Arg Phe Gln Pro Gln Asn Pro
 50 55 60
 Leu Ser Gly Ala Gln Gln Phe Val Ala Lys Asp Pro Gln Asp Asp Asp
 65 70 75 80
 Asp Leu Lys Leu Cys Ser His Thr Met Met Leu Pro Thr Arg Gly Gln
 85 90 95
 Leu Glu Gly Arg Met Ile Val Thr Ala Tyr Glu His Gly Leu Asp Asn
 100 105 110
 Val Thr Glu Glu Ala Val Ser Ala Val Val Tyr Ala Val Glu Asn His
 115 120 125
 Leu Lys Asp Ile Leu Thr Ser Val Val Ser Arg Arg Lys Ala Tyr Arg
 130 135 140
 Leu Arg Asp Gly His Phe Lys Tyr Ala Phe Gly Ser Asn Val Thr Pro
 145 150 155 160
 Gln Pro Tyr Leu Lys Asn Ser Val Val Ala Tyr Asn Asn Leu Ile Glu
 165 170 175
 Ser Pro Pro Ala Phe Thr Ala Pro Cys Ala Gly Gln Asn Pro Ala Ser
 180 185 190
 His Pro Pro Pro Asp Asp Ala Glu Gln Gln Ala Ala Leu Leu Leu Ala
 195 200 205
 Cys Ser Gly Asp Thr Leu Pro Ala Ser Leu Pro Pro Val Asn Met Tyr
 210 215 220

Asp Leu Phe Glu Ala Leu Gln Val His Arg Glu Val Ile Pro Thr His
 225 230 235 240

Thr Val Tyr Ala Leu Asn Ile Glu Arg Ile Ile Thr Lys Leu Trp His
 245 250 255

Pro Asn His Glu Glu Leu Gln Gln Asp Lys Val His Arg Gln Arg Leu
 260 265 270

Ala Ala Lys Glu Gly Leu Leu Leu Cys Xaa
 275 280

<210> 211

<211> 48

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (48)

<223> Xaa equals stop translation

<400> 211

Met Pro Lys Thr Cys Leu Pro Ile Leu Cys Leu Pro Leu Thr Gln Ala
 1 5 10 15

Val Val Leu Ala Gln Leu Asn Asn Phe Ser Ser Leu Asn Ile Phe Ile
 20 25 30

Phe Lys Ile Lys Asn Lys Met Tyr Tyr Ile Trp Ile Tyr Asp Lys Xaa
 35 40 45

<210> 212

<211> 59

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (59)

<223> Xaa equals stop translation

<400> 212

Met Trp Pro Cys Cys Leu Asp Ser Leu Leu Phe Gly Phe Trp Leu Trp
 1 5 10 15

Ala Gln Gly Ile Thr Leu Leu Ser Glu Asp Ser Ile Arg Ile Val Cys
 20 25 30

Ser Ser Cys Glu Pro Glu Val Leu His Val Pro Thr Pro Val Tyr Arg
 35 40 45

Pro Cys Pro Ser His Ser Pro Leu Thr Phe Xaa
 50 55

<210> 213
 <211> 43
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (43)
 <223> Xaa equals stop translation

<400> 213
 Met Ala Leu Gln Ser Ile Pro Ser Phe Thr Leu Leu Ile Ser Phe Phe
 1 5 10 15
 Leu Ser Thr Gln Cys Leu Arg Cys Val Tyr Asn Tyr Glu Cys Ile Leu
 20 25 30
 Phe Met Ala Phe Asn Cys Arg Met Val Phe Xaa
 35 40

<210> 214
 <211> 53
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (53)
 <223> Xaa equals stop translation

<400> 214
 Met Pro Ala Val Ser Ala Phe Phe Ser Leu Ala Ala Leu Ala Glu Val
 1 5 10 15
 Ala Ala Met Glu Asn Val His Arg Gly Gln Arg Ser Thr Pro Leu Thr
 20 25 30
 His Asp Gly Gln Pro Lys Glu Met Pro Gln Ala Pro Val Leu Ile Ser
 35 40 45
 Cys Ala Asp Gln Xaa
 50

<210> 215
 <211> 68
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE

<222> (68)

<223> Xaa equals stop translation

<400> 215

Met Cys Thr Gln Ile Leu Val Phe Met Leu Leu Ile Lys Cys Ile Phe
 1 5 10 15

Ser Ile Asn Thr His Pro Ile Met Pro Tyr Leu Tyr Met Lys Asn Lys
 20 25 30

Val Thr Met Leu Tyr Cys Tyr Val Leu Lys Phe Lys Ser Leu Phe Glu
 35 40 45

Lys Pro Ser Asn Trp Cys Phe His Tyr Ile Met Ile His Leu Asp Lys
 50 55 60

Thr Pro Asn Xaa
 65

<210> 216

<211> 57

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (57)

<223> Xaa equals stop translation

<400> 216

Met Leu Phe Val Ser Leu Leu Val Met Trp Asn Leu Phe Leu Ser Ser
 1 5 10 15

Asp Phe Leu Phe Leu Trp Ser Val Leu Gly Tyr Tyr Met Lys Val Arg
 20 25 30

Leu Pro Gln Ser Pro Arg Glu Ala His Cys Val Leu Leu Ile Asp Leu
 35 40 45

Lys Met Ile Glu Ser Leu Gly Gly Xaa
 50 55

<210> 217

<211> 56

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (56)

<223> Xaa equals stop translation

<400> 217

Met Glu Gln Leu Leu Ala Ala Val Val Phe Phe Ser Ile Phe Phe Leu
 1 5 10 15

Asn Leu Leu Ala Leu Lys Met Asn Lys Val Tyr Arg Cys Ile Cys Leu
 20 25 30
 Leu Phe Ser Lys Asn Met His Thr Asn Val Cys Phe Tyr Lys Ser Asn
 35 40 45
 Thr His Val Ile Ile Cys Met Xaa
 50 55

<210> 218
 <211> 58
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (58)
 <223> Xaa equals stop translation

<400> 218
 Met Cys Trp Lys Pro Lys Cys Ile Leu Leu Leu Ser Phe Val Phe Gln
 1 5 10 15
 Cys Val Ala Ser Ser Thr Phe Asp Pro Leu Gly Ser Glu Arg Pro Trp
 20 25 30
 Ser Gln Pro Gln Cys Pro Ile Ser Phe Pro Leu Leu Ile Thr Gly Cys
 35 40 45
 Cys Trp Phe Ser Met Ser Arg Val Ser Xaa
 50 55

<210> 219
 <211> 59
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (59)
 <223> Xaa equals stop translation

<400> 219
 Met Arg Thr Phe Leu Thr Phe Val Ile Leu Lys Val Ile Leu Ile Phe
 1 5 10 15
 Leu Ser Ser Cys Ala Ser Phe Thr Arg Asn Leu Leu Thr Trp Pro Asn
 20 25 30
 Asp Val Ser Thr Glu Gln Phe Glu Thr Arg Pro Phe Gly Ser Glu Leu
 35 40 45
 Leu Gln Thr Val Ile Asn Val Ser Arg Thr Xaa
 50 55

<210> 220
 <211> 45
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (45)
 <223> Xaa equals stop translation

<400> 220
 Met Arg Phe Phe Phe Gln Ala Tyr Ser Gln Ile Cys Val Gln Asn Phe
 1 5 10 15

Leu Thr Phe Leu Leu Cys Ile Ile Ile Glu Phe Ile Ala Ala Asp Phe
 20 25 30

Tyr Asn Asp Ser Cys Cys His Val Ser Leu Asn Asn Xaa
 35 40 45

<210> 221
 <211> 45
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (41)
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>
 <221> SITE
 <222> (45)
 <223> Xaa equals stop translation

<400> 221
 Met Ile Leu Phe Asp Leu Thr Phe Phe Leu Phe Ala Pro Arg Ile Leu
 1 5 10 15

Ala Ser Gly Ala Cys Ser Cys Ser Ile Tyr Pro Lys Ile Thr Leu Pro
 20 25 30

Thr Lys Tyr Phe Ala Phe Ile Ile Xaa Thr Ser Phe Xaa
 35 40 45

<210> 222
 <211> 52
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (52)

132

<223> Xaa equals stop translation

<400> 222

Met Asp Gly Leu Ile Met Cys Leu Ile Ile Phe Gln Ile Val Asn Phe
 1 5 10 15

Trp Leu Pro Cys Ile Ile Leu Leu Gly Ile Leu Asn Pro Thr Tyr Lys
 20 25 30

Asn Tyr Val Met Val Ser Thr Lys Cys Trp Met Lys Arg Thr Tyr Glu
 35 40 45

His Met Ser Xaa
 50

<210> 223

<211> 73

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (73)

<223> Xaa equals stop translation

<400> 223

Met Thr Phe Leu Phe Phe Phe Leu Phe Ser Arg Ile Leu Cys Ile Lys
 1 5 10 15

Asn Leu Asp Leu Leu Thr Trp Lys Arg Ser Asn Pro Val Ile Ala Lys
 20 25 30

His Leu Tyr Cys Arg Gly His Ile Thr Lys Lys Ser Lys Gly Pro Ala
 35 40 45

Gln Trp Thr Ile Tyr Phe Ser Asp Val Gln Tyr Lys Ile Ser Leu Pro
 50 55 60

Leu Lys Thr Leu Glu Ser Pro Phe Xaa
 65 70

<210> 224

<211> 71

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (71)

<223> Xaa equals stop translation

<400> 224

Met Leu Phe Trp Lys Phe Gly Ser Phe Leu Phe Phe Cys Leu Pro Leu
 1 5 10 15

133

Thr Leu Phe Cys Ile Leu Asn Glu Arg Gly Ile Met His Leu Glu Gly
 20 25 30

Gly Thr Leu Leu Asn Ser Leu Ser His Val Arg His Tyr Leu Arg Leu
 35 40 45

Arg Leu Ser Cys Phe Glu Lys Ile Pro Leu His Arg Ser Ile Phe Ile
 50 55 60

Phe Leu Leu Leu Leu Xaa
 65 70

<210> 225

<211> 58

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (58)

<223> Xaa equals stop translation

<400> 225

Met Ala Gly Cys Cys Leu Lys Leu Phe Gly Val Leu Ser Leu Cys Phe
 1 5 10 15

Leu Cys Gly Leu Ile Ser Ile Glu Arg Val Ile Cys Asn Pro Val Ser
 20 25 30

Ala Asp Phe Gln Val Ser Thr Phe Cys Gln Arg His Cys Leu Leu Arg
 35 40 45

Ser Lys Val Met Phe Pro Ile Arg Gly Xaa
 50 55

<210> 226

<211> 59

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (59)

<223> Xaa equals stop translation

<400> 226

Met Arg Ile Ser Arg Cys Asn Ile Ser Leu Glu Ile Val Ser Pro Ser
 1 5 10 15

Ile Leu Leu Thr Phe Leu Asp Leu Ile Ile Leu Leu Trp Ala Leu Ala
 20 25 30

Ser Cys Tyr Arg Arg Phe Thr Ser Phe Pro Ala Leu Asn Leu Pro Asp
 35 40 45

134

Val Asn Ser Thr Leu His Tyr Leu Gln Gln Xaa
 50 55

<210> 227
 <211> 43
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (43)
 <223> Xaa equals stop translation

<400> 227
 Met Val Ala Pro Leu His Leu Phe Ile Pro Phe Ser Trp Leu Val Arg
 1 5 10 15

Thr Ile Gly Gln Leu Leu Ser Pro Val Gly Lys Ala Leu Ser His Arg
 20 25 30

Ser Asn Gln Met Met Pro Arg Ser Trp Gly Xaa
 35 40

<210> 228
 <211> 41
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (41)
 <223> Xaa equals stop translation

<400> 228
 Met Arg Thr Ser Leu Phe Phe Phe Phe Phe Lys Asn Ile Leu Val Leu
 1 5 10 15

Cys Gly Thr Leu Leu Ile Ser Arg Ser Ser His Ser Gln Ser Ala Pro
 20 25 30

Arg Gly Cys Trp Trp Pro His Lys Xaa
 35 40

<210> 229
 <211> 42
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (42)
 <223> Xaa equals stop translation

<400> 229

135

Met Leu Trp Lys Tyr Phe Leu Ser Leu Phe Leu Pro Trp Tyr Leu Tyr
 1 5 10 15

Cys Phe Phe Asn Asn Asn Ile Met Phe Tyr Ser Leu His Ser Val Pro
 20 25 30

Met Phe Ile Gln Pro Phe Leu Leu Trp Xaa
 35 40

<210> 230

<211> 165

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (165)

<223> Xaa equals stop translation

<400> 230

Met Ser Thr Arg Arg Leu Gly Val Ala Val Ala Val Leu Gly Gly Phe
 1 5 10 15

Leu Tyr Ala Val Gly Gly Ser Asp Gly Thr Ser Pro Leu Asn Thr Val
 20 25 30

Glu Arg Tyr Asn Pro Gln Glu Asn Arg Trp His Thr Ile Ala Pro Met
 35 40 45

Gly Thr Arg Arg Lys His Leu Gly Cys Ala Val Tyr Gln Asp Met Ile
 50 55 60

Tyr Ala Val Gly Gly Arg Asp Asp Thr Thr Glu Leu Ser Ser Ala Glu
 65 70 75 80

Arg Tyr Asn Pro Arg Thr Asn Gln Trp Ser Pro Val Val Ala Met Thr
 85 90 95

Ser Arg Arg Ser Gly Val Gly Leu Ala Val Val Asn Gly Gln Leu Met
 100 105 110

Ala Val Gly Gly Phe Asp Gly Thr Thr Tyr Leu Lys Thr Ile Glu Val
 115 120 125

Phe Asp Pro Asp Ala Asn Thr Trp Arg Leu Tyr Gly Gly Met Asn Tyr
 130 135 140

Arg Arg Leu Gly Gly Gly Val Gly Val Ile Lys Met Thr His Cys Glu
 145 150 155 160

Ser His Ile Trp Xaa
 165

<210> 231

<211> 52

136

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (52)

<223> Xaa equals stop translation

<400> 231

Met Ala Cys Leu Ile Arg Phe Pro Ala Ile Gly Ser Leu Pro Tyr Ser
 1 5 10 15

Thr Trp Pro Phe Phe Phe Ile Phe Leu Phe Phe Ser Cys Leu Thr
 20 25 30

Phe Ile Pro Phe Ser Pro Leu Ser Ser Phe Cys Glu Pro Tyr Pro Arg
 35 40 45

Lys Glu Pro Xaa
 50

<210> 232

<211> 130

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (130)

<223> Xaa equals stop translation

<400> 232

Met Phe Leu Leu Asn Phe Arg Tyr Ile Met Arg Phe Phe Phe Trp Pro
 1 5 10 15

Met Leu Gln Ala Lys Leu Met Ser Phe His Phe Leu Lys Pro Ile Ile
 20 25 30

Phe Met Asn Ser Leu Ile Leu Cys Leu Lys Gln Ser Cys Ser Cys Glu
 35 40 45

Val Glu Ile Ser Leu Leu Pro Leu Ser Gln Gln Thr His Arg Thr Asp
 50 55 60

Leu Gly Phe Ser His Ser Gly Ser Gln Asn Glu Pro Phe Leu Asn Leu
 65 70 75 80

Asp Lys Arg Ala Ala Glu Ala His Cys Ala Val Met Val Leu Cys Leu
 85 90 95

Leu Gly Arg Asp Leu Lys Ala Arg Arg Ser Arg Glu Gly Pro Ala Leu
 100 105 110

Cys Ser Ser Gln Val Val Ile Cys Ile Leu Lys Leu Ala Arg Lys Arg
 115 120 125

Phe Xaa
130

<210> 233
<211> 55
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (55)
<223> Xaa equals stop translation

<400> 233
Met Glu Phe Lys Leu Val Arg Lys Ile Gln Ile Ala Ile Leu Ile Phe
1 5 10 15
Tyr Leu Tyr Leu Val Ala Val Ala Phe Lys Asn Lys Phe Ser Tyr Lys
20 25 30
Ser Phe Gln Phe Phe Gly Leu Glu Ser Ile Phe Gln Asn Lys Lys Leu
35 40 45
Lys Lys Glu Tyr Leu Met Xaa
50 55

<210> 234
<211> 363
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (307)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (363)
<223> Xaa equals stop translation

<400> 234
Met Arg Thr Leu Phe Asn Leu Leu Trp Leu Ala Leu Ala Cys Ser Pro
1 5 10 15
Val His Thr Thr Leu Ser Lys Ser Asp Ala Lys Lys Ala Ala Ser Lys
20 25 30
Thr Leu Leu Glu Lys Ser Gln Phe Ser Asp Lys Pro Val Gln Asp Arg
35 40 45
Gly Leu Val Val Thr Asp Leu Lys Ala Glu Ser Val Val Leu Glu His
50 55 60
Arg Ser Tyr Cys Ser Ala Lys Ala Arg Asp Arg His Phe Ala Gly Asp

PCT/US98/27059

SUBSTITUTE SHEET (RULE 26)

139

<210> 235
 <211> 29
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (29)
 <223> Xaa equals stop translation

<400> 235
 Met Cys Met Cys Val Leu Leu Cys Val Phe Leu Ile Cys Lys Tyr Ser
 1 5 10 15
 Lys Ser Phe Leu Ile Leu Arg Leu Lys Phe Ser Cys Xaa
 20 25

<210> 236
 <211> 67
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (67)
 <223> Xaa equals stop translation

<400> 236
 Met Gly Asn Ala Cys Ile Pro Leu Lys Arg Ile Ala Tyr Phe Leu Cys
 1 5 10 15
 Leu Leu Ser Ala Leu Leu Leu Thr Glu Gly Lys Lys Pro Ala Asn Gln
 20 25 30
 Asn Ala Leu Pro Cys Val Leu Val Pro Lys Ile Met Leu Tyr Val Arg
 35 40 45
 Met Pro Asp Pro Phe His Ala Pro Phe Leu Leu Met Leu Ser His Tyr
 50 55 60
 Pro Leu Xaa
 65

<210> 237
 <211> 114
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (114)
 <223> Xaa equals stop translation

<400> 237
 Met Ile Leu Ser Leu Leu Phe Ser Leu Gly Gly Pro Leu Gly Trp Gly

1 5 10 15
 Leu Leu Gly Ala Trp Ala Gln Ala Ser Ser Thr Ser Leu Ser Asp Leu
 20 25 30
 Gln Ser Ser Arg Thr Pro Gly Val Trp Lys Ala Glu Ala Glu Asp Thr
 35 40 45
 Ser Lys Asp Pro Val Gly Arg Asn Trp Cys Pro Tyr Pro Met Ser Lys
 50 55 60
 Leu Val Thr Leu Leu Ala Leu Cys Lys Thr Glu Lys Phe Leu Ile His
 65 70 75 80
 Ser Gln Gln Pro Cys Pro Gln Glu Leu Gln Thr Ala Arg Lys Ser Lys
 85 90 95
 Ser Cys Thr Ala Trp Pro Thr Ser Gln Cys Thr Arg Ser Ser Arg Arg
 100 105 110
 Cys Xaa

<210> 238
 <211> 106
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (106)
 <223> Xaa equals stop translation

<400> 238
 Met Ala Ile His Lys Ala Leu Val Met Cys Leu Gly Leu Pro Leu Phe
 1 5 10 15
 Leu Phe Pro Gly Ala Trp Ala Gln Gly His Val Pro Pro Gly Cys Ser
 20 25 30
 Gln Gly Leu Asn Pro Leu Tyr Tyr Asn Leu Cys Asp Arg Ser Gly Ala
 35 40 45
 Trp Gly Ile Val Leu Glu Ala Val Ala Gly Ala Gly Ile Val Thr Thr
 50 55 60
 Phe Val Leu Thr Ile Ile Leu Val Ala Ser Leu Pro Phe Val Gln Asp
 65 70 75 80
 Thr Lys Lys Arg Ser Leu Leu Gly Thr Gln Leu Arg Gly Arg Cys His
 85 90 95
 His Thr Ala Gly Thr Met Gly Ser Cys Xaa
 100 105

141

<210> 239
 <211> 15
 <212> PRT
 <213> Homo sapiens

<400> 239
 Gly Leu Gly Pro Ala Gln Val Ala Leu Ser Leu Gln Gly Pro Ala
 1 5 10 15

<210> 240
 <211> 82
 <212> PRT
 <213> Homo sapiens

<400> 240
 Ser Ser Trp Met Ala Gly Thr Gln Pro Arg Thr Ser Trp Trp Glu Met
 1 5 10 15

Ser Ser Ala Lys Pro Cys Pro Thr Gly Thr Leu Arg Ser Asn Thr Ser
 20 25 30

Ser His Pro Gln Cys Thr Gly Pro Pro Thr Thr His Pro Met Leu Val
 35 40 45

Gly Glu Asp Met Ser Cys Pro Glu Pro Gln Cys Gly Ala Ser Arg Leu
 50 55 60

Ser Trp Lys Met Leu Asn Ser Ser Pro Leu Met Met Ser Leu Trp Val
 65 70 75 80

Cys Ala

<210> 241
 <211> 23
 <212> PRT
 <213> Homo sapiens

<400> 241
 Gln Pro Arg Thr Ser Trp Trp Glu Met Ser Ser Ala Lys Pro Cys Pro
 1 5 10 15

Thr Gly Thr Leu Arg Ser Asn
 20

<210> 242
 <211> 23
 <212> PRT
 <213> Homo sapiens

<400> 242
 Met Ser Cys Pro Glu Pro Gln Cys Gly Ala Ser Arg Leu Ser Trp Lys
 1 5 10 15

142

Met Leu Asn Ser Ser Pro Leu
20

<210> 243
<211> 98
<212> PRT
<213> Homo sapiens

<400> 243
Trp Val Ala Leu Tyr Ile Glu Gly Gly Met Lys Tyr Leu Thr Leu Val
1 5 10 15
Phe Leu Leu Gly Arg Ala Trp Arg Met Thr Ser Pro Thr Arg Arg Ser
20 25 30
Trp Ala Gly Ser Gln Pro Ser Arg Asn Ser Asn Thr Leu Gly Thr Trp
35 40 45
Thr Lys Thr Ser Ser Ser Pro Phe Ser Met Lys Trp Ala Trp Gly Gln
50 55 60
Ala Ala Thr Thr Gln Arg Cys Arg Cys Ser Ser Leu Ser Val Arg Leu
65 70 75 80
Lys Lys Ser Ser Val Lys Ser His Trp Arg Met Ser Ser Asn Ser Leu
85 90 95
Leu Ser

<210> 244
<211> 20
<212> PRT
<213> Homo sapiens

<400> 244
Gly Gly Met Lys Tyr Leu Thr Leu Val Phe Leu Leu Gly Arg Ala Trp
1 5 10 15
Arg Met Thr Ser
20

<210> 245
<211> 25
<212> PRT
<213> Homo sapiens

<400> 245
Ser Gln Pro Ser Arg Asn Ser Asn Thr Leu Gly Thr Trp Thr Lys Thr
1 5 10 15
Ser Ser Ser Pro Phe Ser Met Lys Trp
20 25

<210> 246
<211> 26
<212> PRT
<213> Homo sapiens

<400> 246
Thr Thr Gln Arg Cys Arg Cys Ser Ser Leu Ser Val Arg Leu Lys Lys
1 5 10 15

Ser Ser Val Lys Ser His Trp Arg Met Ser
20 25

<210> 247
<211> 223
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (13)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (14)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (15)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (27)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (108)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (113)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (117)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (121)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (122)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (125)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (129)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (130)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 247

Ala	Ser	Thr	Leu	Ala	Gln	Thr	Thr	Gly	Thr	Cys	Lys	Xaa	Xaa	Xaa	Ser
1				5					10						15

Ser	Arg	Arg	Ala	Arg	Ser	Arg	Thr	Gln	Arg	Xaa	Phe	Gln	Leu	Arg	Pro
			20					25						30	

Asp	Lys	Arg	Ser	Ala	Pro	Ser	Leu	Leu	Gln	Phe	Ile	Gln	Ala	Gln	Glu
	35						40					45			

Glu	Leu	Ser	Lys	Glu	Asn	Thr	Gly	Arg	Gln	Leu	Ala	Ala	Arg	Glu	Ala
	50					55					60				

Val	Leu	Ala	Leu	Glu	Gly	Ser	Thr	Gln	Leu	Thr	Gly	Pro	Val	Thr	Gln
	65					70				75					80

Val	Ala	Ala	Ser	Lys	Thr	His	Cys	Ser	Gly	Met	Ala	Leu	Thr	Ala	Ser
				85					90					95	

Pro	Val	Pro	Val	Leu	Gly	Ala	Ala	Pro	Ala	Lys	Xaa	Pro	Thr	Gln	Asn
			100					105						110	

Xaa	Pro	Gly	Gln	Xaa	Gly	Arg	Ala	Xaa	Xaa	Lys	Val	Xaa	Thr	Ser	Trp
		115				120						125			

Xaa	Xaa	Val	Ala	Thr	Lys	Val	Leu	His	Gly	Leu	Glu	Val	Ser	Thr	His
		130				135					140				

Leu	Gly	Lys	Arg	Lys	Leu	Ser	Gly	Arg	Ser	Trp	Leu	Pro	Gly	Pro	Ala
	145				150					155					160

Leu	His	Ala	Thr	Pro	Ser	Gln	Ser	His	Thr	Gln	Thr	Gly	Ser	Gln	Ile
				165					170					175	

Val	His	Pro	Pro	Gln	Gly	Glu	Val	Arg	Glu	Val	Gly	Arg	Gly	Arg	Gly
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

145

180 185 190

Gln Pro Pro Ala Gln Pro Val His Ala His Pro Ser Gln Gln His Pro
 195 200 205

Ser Pro Ala His Leu Ala Gly Leu Ser Leu Trp Thr Gly Thr Ala
 210 215 220

<210> 248
 <211> 140
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (12)
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>
 <221> SITE
 <222> (59)
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>
 <221> SITE
 <222> (60)
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>
 <221> SITE
 <222> (80)
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>
 <221> SITE
 <222> (82)
 <223> Xaa equals any of the naturally occurring L-amino acids

<400> 248

Ala Met Leu Glu Thr Trp Arg Pro Gly Pro Ser Xaa Gly Glu Leu Ala
 1 5 10 15

Thr Asn Ser Gly Gln Arg Ala Ser Gln Asp Ser Gln His Ser Pro Pro
 20 25 30

His Val Arg Ala His Leu Leu Ile Ser Pro Leu Pro Ala Phe Pro Ser
 35 40 45

Met Gly Gly Pro Ala Gly Arg Ser Ala Pro Xaa Xaa Leu Thr Glu Thr
 50 55 60

Lys Ser Glu Leu Gln Arg Leu Arg Arg Arg Gln Ala Arg Ala Ser Xaa
 65 70 75 80

Ser Xaa Pro Ala Gly Glu Pro Gly Ala Gly His Ser Asp Ser Phe Asn
 85 90 95

146

Cys Val Pro Thr Asn Gly Gln Pro Leu Arg Ser Cys Ser Leu Ser Lys
 100 105 110

Leu Arg Arg Ser Phe Leu Lys Arg Thr Gln Gly Asp Ser Trp Leu Pro
 115 120 125

Glu Lys Gln Ser Trp Leu Trp Lys Ala Pro Pro Ser
 130 135 140

<210> 249

<211> 122

<212> PRT

<213> Homo sapiens

<400> 249

Ser His Gln Ser His Leu Ile Asn Pro Ala Ser Ser Ala Lys Gly Ser
 1 5 10 15

Trp Ala Gln Leu Lys Ala Gln Pro Pro Ala His Val Leu Gly Gly Thr
 20 25 30

Gly Gln Glu Gly Pro Pro Pro Thr Ala Asp Gln Pro Glu Ser Pro Gly
 35 40 45

Trp Asp Pro Ser Ser Phe Thr Asn Gly Ser Ser Gly Pro Arg Ala Leu
 50 55 60

Pro Thr Ser Val His Pro Thr Leu Gln Gln Gly Ala Pro Cys Arg Arg
 65 70 75 80

Asn Trp Ala Pro Cys Arg Gly Leu Val Glu Thr Arg Met Leu Arg Arg
 85 90 95

Gln Leu Pro His Gly Thr Ser Lys Arg Asp Leu Gly Trp Ala Ser Leu
 100 105 110

Gln Arg Gly Ser Pro Gln Glu Thr Pro Gln
 115 120

<210> 250

<211> 35

<212> PRT

<213> Homo sapiens

<400> 250

Arg Pro Asp Lys Arg Ser Ala Pro Ser Leu Leu Gln Phe Ile Gln Ala
 1 5 10 15

Gln Glu Glu Leu Ser Lys Glu Asn Thr Gly Arg Gln Leu Ala Ala Arg
 20 25 30

Glu Ala Val
 35

<210> 251
 <211> 33
 <212> PRT
 <213> Homo sapiens

<400> 251
 Ala Thr Pro Ser Gln Ser His Thr Gln Thr Gly Ser Gln Ile Val His
 1 5 10 15

Pro Pro Gln Gly Glu Val Arg Glu Val Gly Arg Gly Arg Gly Gln Pro
 20 25 30

Pro

<210> 252
 <211> 29
 <212> PRT
 <213> Homo sapiens

<400> 252
 Gln Asp Ser Gln His Ser Pro Pro His Val Arg Ala His Leu Leu Ile
 1 5 10 15

Ser Pro Leu Pro Ala Phe Pro Ser Met Gly Gly Pro Ala
 20 25

<210> 253
 <211> 28
 <212> PRT
 <213> Homo sapiens

<400> 253
 Asp Ser Phe Asn Cys Val Pro Thr Asn Gly Gln Pro Leu Arg Ser Cys
 1 5 10 15

Ser Leu Ser Lys Leu Arg Arg Ser Phe Leu Lys Arg
 20 25

<210> 254
 <211> 25
 <212> PRT
 <213> Homo sapiens

<400> 254
 Lys Gly Ser Trp Ala Gln Leu Lys Ala Gln Pro Pro Ala His Val Leu
 1 5 10 15

Gly Gly Thr Gly Gln Glu Gly Pro Pro
 20 25

<210> 255

<211> 26
 <212> PRT
 <213> Homo sapiens

<400> 255
 Ala Pro Ser Leu Leu Gln Phe Ile Gln Ala Gln Glu Glu Leu Ser Lys
 1 5 10 15
 Glu Asn Thr Gly Arg Gln Leu Ala Ala Arg
 20 25

<210> 256
 <211> 6
 <212> PRT
 <213> Homo sapiens

<400> 256
 Lys Pro Ser His Gln Pro
 1 5

<210> 257
 <211> 21
 <212> PRT
 <213> Homo sapiens

<400> 257
 Cys Ser Tyr Arg Pro Gln Phe Pro Val Asp Pro Arg Val Arg Ala Thr
 1 5 10 15

Cys Ile Val Phe Asn
 20

<210> 258
 <211> 128
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (46)
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>
 <221> SITE
 <222> (60)
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>
 <221> SITE
 <222> (66)
 <223> Xaa equals any of the naturally occurring L-amino acids

<400> 258
 Gly Thr Glu Asn Leu Leu Ala Pro Glu Arg Thr Ile Leu Ser Arg Ala

WO 99/31117

PCT/US98/27059

1 5 10 15 149

Gln Met Gly Lys Cys Met Ala Thr Pro Ala Pro Cys Val Arg Ser Ser
20 25 30

Ser Lys Gln Lys Lys Lys Lys Arg Lys Arg Arg Lys Val Xaa Gln Glu
35 40 45

Thr Lys Asp Asn Leu Arg Val Gln Leu Pro Leu Xaa Ser Cys Val Val
50 55 60

Asn Xaa Ala Asn Pro Gly Lys Thr Asp Gly Phe Phe Ala Pro Glu Arg
65 70 75 80

Met Thr Pro Ser Arg Ala Gln Met Glu Lys Cys Met Ala Thr Pro Ala
85 90 95

Pro Cys Val Arg Pro Ser Phe Asn Lys Lys Lys Glu Gln Glu Gln Arg
100 105 110

Leu Lys Glu Lys Leu Gln Arg Lys Ser Ala Val Asn Phe Gly Thr Lys
115 120 125

<210> 259
<211> 26
<212> PRT
<213> Homo sapiens

<400> 259
Leu Leu Ala Pro Glu Arg Thr Ile Leu Ser Arg Ala Gln Met Gly Lys
1 5 10 15
Cys Met Ala Thr Pro Ala Pro Cys Val Arg
20 25

<210> 260
<211> 24
<212> PRT
<213> Homo sapiens

<400> 260
Pro Gly Lys Thr Asp Gly Phe Phe Ala Pro Glu Arg Met Thr Pro Ser
1 5 10 15

Arg Ala Gln Met Glu Lys Cys Met
20

<210> 261
<211> 17
<212> PRT
<213> Homo sapiens

SUBSTITUTE SHEET (RULE 26)

<400> 261

Glu Gln Arg Leu Lys Glu Lys Leu Gln Arg Lys Ser Ala Val Asn Phe
 1 5 10 15

Gly

<210> 262

<211> 186

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (42)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (68)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (69)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 262

Lys Thr Leu Leu Glu Asn Phe Ser Thr Gln Gly Thr Phe Val Ala Met
 1 5 10 15

His Pro Ala Val Arg Ala Thr Asp Trp Ile Thr Leu Pro Cys Thr Lys
 20 25 30

Lys Pro Ser Ile Ser His Leu Phe Phe Xaa Phe Leu Ala Lys Ile Leu
 35 40 45

Phe Ser Ile Ser Ser Asn Ser Ser Phe Thr Leu Ser Leu Gly Ile Phe
 50 55 60

Ser Phe Phe Xaa Xaa Gln Leu Ser Thr His Cys Thr Leu Ile Ala Met
 65 70 75 80

Arg Leu Pro Ile Arg Thr Lys Asn Arg Ile Ile Phe Pro Cys Ala Ser
 85 90 95

Lys Ser Ser Ile Ser Asn Lys Gly Pro Lys Ser Thr Ala Tyr Ile Leu
 100 105 110

Leu Trp Ile Thr Ala Leu Thr Phe Pro Phe Thr Phe Tyr Thr Asn Leu
 115 120 125

Gly Pro Gly Phe Arg Ile Leu Ser Thr Gln Cys Thr Ser Val Val Ile
 130 135 140

151 .

Cys Phe Pro Ile Cys Ala Thr Asn Ser Phe Ile Ile Ile Arg Thr Asp
 145 150 155 160

Lys Ile Pro Ile Ser Phe Ser Phe Phe Lys Ile Ile Thr Ile Gln Leu
 165 170 175

Cys Trp Gly Ser Ser Leu Gly Ser Ser Cys
 180 185

<210> 263
 <211> 22
 <212> PRT
 <213> Homo sapiens

<400> 263
 Met His Pro Ala Val Arg Ala Thr Asp Trp Ile Thr Leu Pro Cys Thr
 1 5 10 15

Lys Lys Pro Ser Ile Ser
 20

<210> 264
 <211> 17
 <212> PRT
 <213> Homo sapiens

<400> 264
 Leu Ile Ala Met Arg Leu Pro Ile Arg Thr Lys Asn Arg Ile Ile Phe
 1 5 10 15

Pro

<210> 265
 <211> 26
 <212> PRT
 <213> Homo sapiens

<400> 265
 Ser Ser Ile Ser Asn Lys Gly Pro Lys Ser Thr Ala Tyr Ile Leu Leu
 1 5 10 15

Trp Ile Thr Ala Leu Thr Phe Pro Phe Thr
 20 25

<210> 266
 <211> 23
 <212> PRT
 <213> Homo sapiens

<400> 266
 Ile Ile Ile Arg Thr Asp Lys Ile Pro Ile Ser Phe Ser Phe Phe Lys
 1 5 10 15

Ile Ile Thr Ile Gln Leu Cys
20

<210> 267
<211> 165
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (147)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (153)
<223> Xaa equals any of the naturally occurring L-amino acids

<400> 267
Asn Asp Gly Gln Cys Leu Ala Tyr Asn Thr Thr His Tyr Arg Glu Arg
1 5 10 15
Ala Met Thr Ser His Ala Arg Val Ser Leu Gly Pro Ser Arg Asp Pro
20 25 30
Leu Glu Arg Pro Pro Phe Phe Phe Phe Phe Phe Phe Phe Phe Phe
35 40 45
Lys Phe Glu His Thr Gly Thr His Gly Thr Leu Val Ser Met His Phe
50 55 60
Ala Ile Trp Ala Thr Asp Arg Ile Met Leu Pro Gly Ala Tyr Lys Cys
65 70 75 80
Ser Ile Pro His Leu Val Pro Lys Phe Thr Ala Asp Phe Leu Cys Ser
85 90 95
Phe Ser Phe Ser Leu Cys Ser Cys Ser Phe Phe Leu Leu Lys Glu Gly
100 105 110
Leu Thr His Gly Ala Gly Val Ala Met His Phe Ser Ile Trp Ala Leu
115 120 125
Asp Gly Val Ile Leu Ser Gly Ala Lys Lys Pro Ser Val Phe Pro Gly
130 135 140
Phe Ala Xaa Phe Thr Thr Gln Leu Xaa Lys Gly Ser Cys Thr Leu Arg
145 150 155 160
Leu Ser Phe Val Ser
165

<210> 268
<211> 22

<212> PRT

<213> Homo sapiens

<400> 268

Cys Leu Ala Tyr Asn Thr Thr His Tyr Arg Glu Arg Ala Met Thr Ser
 1 5 10 15

His Ala Arg Val Ser Leu
 20

<210> 269

<211> 31

<212> PRT

<213> Homo sapiens

<400> 269

Gly Thr Leu Val Ser Met His Phe Ala Ile Trp Ala Thr Asp Arg Ile
 1 5 10 15

Met Leu Pro Gly Ala Tyr Lys Cys Ser Ile Pro His Leu Val Pro
 20 25 30

<210> 270

<211> 24

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (18)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (24)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 270

Gly Val Ile Leu Ser Gly Ala Lys Lys Pro Ser Val Phe Pro Gly Phe
 1 5 10 15

Ala Xaa Phe Thr Thr Gln Leu Xaa
 20

<210> 271

<211> 141

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (26)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>
 <221> SITE
 <222> (38)
 <223> Xaa equals any of the naturally occurring L-amino acids

 <220>
 <221> SITE
 <222> (44)
 <223> Xaa equals any of the naturally occurring L-amino acids

 <220>
 <221> SITE
 <222> (57)
 <223> Xaa equals any of the naturally occurring L-amino acids

 <220>
 <221> SITE
 <222> (58)
 <223> Xaa equals any of the naturally occurring L-amino acids

 <400> 271
 Lys Lys Ala Ser His Met Glu Gln Val Leu Pro Cys Ile Phe Pro Ser
 1 5 10 15
 Gly Pro Trp Met Gly Ser Phe Ser Leu Xaa Gln Lys Ser Arg Pro Phe
 20 25 30
 Phe Leu Asp Leu Arg Xaa Ser Leu His Asn Ser Xaa Lys Glu Ala Val
 35 40 45
 Leu Leu Asp Cys Leu Leu Phe Leu Xaa Xaa Pro Ser Phe Phe Phe Phe
 50 55 60
 Ser Ser Ser Ser Ala Trp Lys Lys Thr Ser His Met Glu Gln Val Leu
 65 70 75 80
 Pro Cys Thr Phe Pro Ser Gly Pro Trp Ile Gly Leu Phe Ser Leu Val
 85 90 95
 Gln Ala Ser Phe Pro Phe Leu Thr Ser Phe Arg Tyr Ser Leu Gln Ser
 100 105 110
 Ser Ala Tyr Glu Val Ala Phe Pro Asp Ser Leu Leu Phe Leu Ala Arg
 115 120 125
 Ala Ser Ala Phe Phe Phe Ser Ser Phe Ser Ala Trp Lys
 130 135 140

<210> 272
 <211> 28
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (15)

155

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (27)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 272

Cys	Ile	Phe	Pro	Ser	Gly	Pro	Trp	Met	Gly	Ser	Phe	Ser	Leu	Xaa	Gln
1				5					10					15	

Lys	Ser	Arg	Pro	Phe	Phe	Leu	Asp	Leu	Arg	Xaa	Ser
			20					25			

<210> 273

<211> 28

<212> PRT

<213> Homo sapiens

<400> 273

Trp	Ile	Gly	Leu	Phe	Ser	Leu	Val	Gln	Ala	Ser	Phe	Pro	Phe	Leu	Thr
1				5					10					15	

Ser	Phe	Arg	Tyr	Ser	Leu	Gln	Ser	Ser	Ala	Tyr	Glu
			20				25				

<210> 274

<211> 79

<212> PRT

<213> Homo sapiens

<400> 274

Asn	Ser	Ala	Val	Asn	Ile	Lys	Ile	Arg	Gln	Arg	Met	Glu	Tyr	Phe	Ser
1				5					10					15	

Val	Pro	Glu	Lys	Met	Thr	Leu	Phe	Val	Val	Gln	Met	Gly	Lys	Cys	Met
			20					25					30		

Ala	Thr	Cys	Val	Pro	Cys	Val	Lys	Pro	Thr	Ser	Lys	Gln	Lys	Met	Lys
		35					40					45			

Lys	Arg	Lys	Arg	Leu	Lys	His	Glu	Leu	Glu	Thr	Lys	Glu	Asn	Leu	Glu
	50					55					60				

Lys	Gln	Pro	His	Met	Gln	Ser	Phe	Ala	Val	Asn	Ile	Glu	Ser	Leu
65					70						75			

<210> 275

<211> 23

<212> PRT

<213> Homo sapiens

<400> 275

Ile	Lys	Ile	Arg	Gln	Arg	Met	Glu	Tyr	Phe	Ser	Val	Pro	Glu	Lys	Met
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

SUBSTITUTE SHEET (RULE 26)

WO 99/31117

PCT/US98/27059

156

1

5

10

15

Thr Leu Phe Val Val Gln Met
20

<210> 276

<211> 25

<212> PRT

<213> Homo sapiens

<400> 276

Val Lys Pro Thr Ser Lys Gln Lys Met Lys Lys Arg Lys Arg Leu Lys
1 5 10 15

His Glu Leu Glu Thr Lys Glu Asn Leu
20 25

<210> 277

<211> 63

<212> PRT

<213> Homo sapiens

<400> 277

Pro Arg Val Arg Gly Thr Val Val Arg Leu Arg Gln His Arg Pro Ser
1 5 10 15

Ala Tyr Ile Leu Val Ser Thr Val Leu Thr Leu Met Val Pro Trp His
20 25 30

Ser Leu Asp Pro Asp Ser Ala Leu Ala Asp Ala Phe Tyr Gln Arg Gly
35 40 45

Tyr Arg Trp Ala Gly Phe Ile Val Ala Ala Gly Ser Ile Cys Ala
50 55 60

<210> 278

<211> 25

<212> PRT

<213> Homo sapiens

<400> 278

Thr Val Val Arg Leu Arg Gln His Arg Pro Ser Ala Tyr Ile Leu Val
1 5 10 15

Ser Thr Val Leu Thr Leu Met Val Pro
20 25

<210> 279

<211> 26

<212> PRT

<213> Homo sapiens

<400> 279

SUBSTITUTE SHEET (RULE 26)

WO 99/31117

PCT/US98/27059

157

Trp His Ser Leu Asp Pro Asp Ser Ala Leu Ala Asp Ala Phe Tyr Gln
1 5 10 15

Arg Gly Tyr Arg Trp Ala Gly Phe Ile Val
20 25

<210> 280

<211> 101

<212> PRT

<213> Homo sapiens

<400> 280

Thr Pro Ser Cys Ser Ala Ser Ser Ser Pro Cys His Ala Leu Ser Met
1 5 10 15

Pro Trp Pro Pro Met Gly Ser Ser Ser Arg Cys Leu Pro Met Cys Thr
20 25 30

Pro Gly His Arg Cys Leu Trp Arg Ala Pro Trp Arg Ser Gly Ser Ser
35 40 45

Arg Pro Ser Trp His Cys Cys Trp Thr Trp Ser Arg Trp Phe Ser Ser
50 55 60

Cys Pro Leu Ala His Ser Trp Pro Thr His Ser Trp Pro Pro Val Ser
65 70 75 80

Leu Cys Cys Ala Ser Arg Ser Leu Pro Arg Pro Ala Pro Gln Ala Gln
85 90 95

Pro Ala Leu Ala Pro
100

<210> 281

<211> 24

<212> PRT

<213> Homo sapiens

<400> 281

Leu Ser Met Pro Trp Pro Pro Met Gly Ser Ser Ser Arg Cys Leu Pro
1 5 10 15

Met Cys Thr Pro Gly His Arg Cys
20

<210> 282

<211> 27

<212> PRT

<213> Homo sapiens

<400> 282

Ala Pro Trp Arg Ser Gly Ser Ser Arg Pro Ser Trp His Cys Cys Trp
1 5 10 15

SUBSTITUTE SHEET (RULE 26)

158

Thr Trp Ser Arg Trp Phe Ser Ser Cys Pro Leu
 20 25

<210> 283
 <211> 22
 <212> PRT
 <213> Homo sapiens

<400> 283
 Thr His Ser Trp Pro Pro Val Ser Leu Cys Cys Ala Ser Arg Ser Leu
 1 5 10 15

Pro Arg Pro Ala Pro Gln
 20

<210> 284
 <211> 60
 <212> PRT
 <213> Homo sapiens

<400> 284
 Ala Tyr Ile Leu Val Ser Thr Val Leu Thr Leu Met Val Pro Trp His
 1 5 10 15

Ser Leu Asp Pro Asp Ser Ala Leu Ala Asp Ala Phe Tyr Gln Arg Gly
 20 25 30

Tyr Arg Trp Ala Gly Phe Ile Val Ala Ala Gly Ser Ile Cys Ala Met
 35 40 45

Asn Thr Val Leu Leu Ser Leu Leu Phe Ser Leu Pro
 50 55 60

<210> 285
 <211> 31
 <212> PRT
 <213> Homo sapiens

<400> 285
 Pro Trp His Ser Leu Asp Pro Asp Ser Ala Leu Ala Asp Ala Phe Tyr
 1 5 10 15

Gln Arg Gly Tyr Arg Trp Ala Gly Phe Ile Val Ala Ala Gly Ser
 20 25 30

<210> 286
 <211> 27
 <212> PRT
 <213> Homo sapiens

<400> 286
 Arg Ile Val Tyr Ala Met Ala Ala Asp Gly Leu Phe Phe Gln Val Phe
 1 5 10 15

Ala His Val His Pro Arg Thr Gln Val Pro Val
 20 25

<210> 287
 <211> 16
 <212> PRT
 <213> Homo sapiens

<400> 287
 Asp Leu Glu Ser Leu Val Gln Phe Leu Ser Leu Gly Thr Leu Leu Ala
 1 5 10 15

<210> 288
 <211> 15
 <212> PRT
 <213> Homo sapiens

<400> 288
 Tyr Thr Phe Val Ala Thr Ser Ile Ile Val Leu Arg Phe Gln Lys
 1 5 10 15

<210> 289
 <211> 31
 <212> PRT
 <213> Homo sapiens

<400> 289
 Leu Thr Lys Gln Gln Ser Ser Phe Ser Asp His Leu Gln Leu Val Gly
 1 5 10 15

Thr Val His Ala Ser Val Pro Glu Pro Gly Glu Leu Lys Pro Ala
 20 25 30

<210> 290
 <211> 50
 <212> PRT
 <213> Homo sapiens

<400> 290
 Leu Arg Pro Tyr Leu Gly Phe Leu Asp Gly Tyr Ser Pro Gly Ala Val
 1 5 10 15

Val Thr Trp Ala Leu Gly Val Met Leu Ala Ser Ala Ile Thr Ile Gly
 20 25 30

Cys Val Leu Val Phe Gly Asn Ser Thr Leu His Leu Pro His Trp Gly
 35 40 45

Tyr Ile

50

<210> 291
<211> 27
<212> PRT
<213> Homo sapiens

<400> 291
Pro Gly Ala Val Val Thr Trp Ala Leu Gly Val Met Leu Ala Ser Ala
1 5 10 15
Ile Thr Ile Gly Cys Val Leu Val Phe Gly Asn
20 25

<210> 292
<211> 53
<212> PRT
<213> Homo sapiens

<400> 292
Gly Ala His Gln Gln Gln Tyr Arg Glu Asp Leu Phe Gln Ile Pro Met
1 5 10 15
Val Pro Leu Ile Pro Ala Leu Ser Ile Val Leu Asn Ile Cys Leu Met
20 25 30
Leu Lys Leu Ser Tyr Leu Thr Trp Val Arg Phe Ser Ile Trp Leu Leu
35 40 45
Met Gly Leu Ala Val
50

<210> 293
<211> 26
<212> PRT
<213> Homo sapiens

<400> 293
Met Val Pro Leu Ile Pro Ala Leu Ser Ile Val Leu Asn Ile Cys Leu
1 5 10 15
Met Leu Lys Leu Ser Tyr Leu Thr Trp Val
20 25

<210> 294
<211> 29
<212> PRT
<213> Homo sapiens

<400> 294
Tyr Phe Gly Tyr Gly Ile Arg His Ser Lys Glu Asn Gln Arg Glu Leu
1 5 10 15

161

Pro Gly Leu Asn Ser Thr His Tyr Val Val Phe Pro Arg
 20 25

<210> 295
 <211> 23
 <212> PRT
 <213> Homo sapiens

<400> 295
 Phe Pro Pro Ser Pro Ala Pro Pro His Ser Leu Pro Leu Arg Ser Trp
 1 5 10 15

Leu Trp Ser Arg Gln Met Gly
 20

<210> 296
 <211> 148
 <212> PRT
 <213> Homo sapiens

<400> 296
 Gly Thr Ser Phe Arg Gly Met Ile Ser Thr Gln Pro Gly Ser Thr Pro
 1 5 10 15

Leu Ala Ser Phe Lys Ile Leu Ala Leu Glu Ser Ala Asp Gly His Gly
 20 25 30

Gly Cys Ser Ala Gly Asn Asp Ile Gly Pro Tyr Gly Glu Arg Asp Asp
 35 40 45

Gln Gln Val Phe Ile Gln Lys Val Val Pro Ser Ala Ser Gln Leu Phe
 50 55 60

Val Arg Leu Ser Ser Thr Gly Gln Arg Val Cys Ser Val Arg Ser Val
 65 70 75 80

Asp Gly Ser Pro Thr Thr Ala Phe Thr Val Leu Glu Cys Glu Gly Ser
 85 90 95

Pro Ala Ala Arg Leu Ser Ala Pro Ala Leu Pro Ala His Trp Pro Gly
 100 105 110

Gln Arg Gln Leu Gly His Val Gly Pro Asn His Arg His Gly Arg Pro
 115 120 125

Arg Pro Gly Pro Cys Arg Trp Pro Asp Gly Ala Arg Ala Asp Gly Thr
 130 135 140

Ala Gly Thr Leu
 145

<210> 297
 <211> 29
 <212> PRT

162

<213> Homo sapiens

<400> 297

Pro Gly Ser Thr Pro Leu Ala Ser Phe Lys Ile Leu Ala Leu Glu Ser
1 5 10 15

Ala Asp Gly His Gly Gly Cys Ser Ala Gly Asn Asp Ile
20 25

<210> 298

<211> 24

<212> PRT

<213> Homo sapiens

<400> 298

Gly Glu Arg Asp Asp Gln Gln Val Phe Ile Gln Lys Val Val Pro Ser
1 5 10 15

Ala Ser Gln Leu Phe Val Arg Leu
20

<210> 299

<211> 25

<212> PRT

<213> Homo sapiens

<400> 299

Arg Ser Val Asp Gly Ser Pro Thr Thr Ala Phe Thr Val Leu Glu Cys
1 5 10 15

Glu Gly Ser Pro Ala Ala Arg Leu Ser
20 25

<210> 300

<211> 26

<212> PRT

<213> Homo sapiens

<400> 300

Pro Ala Leu Pro Ala His Trp Pro Gly Gln Arg Gln Leu Gly His Val
1 5 10 15

Gly Pro Asn His Arg His Gly Arg Pro Arg
20 25

<210> 301

<211> 168

<212> PRT

<213> Homo sapiens

<400> 301

Pro Phe Ile Pro Arg Arg Pro Trp Pro Glu Pro Gly Val Pro Thr Gly
1 5 10 15

Ile Arg Glu Ala Pro Glu Ser Pro Arg Thr Arg Ala Ser Gln Gly Ile
 20 25 30

Met Ala Ala Ala Leu Phe Lys Lys Glu Val Ser Leu Ser Phe Ile Leu
 35 40 45

Gly Gly Val Arg Gly Val Pro Arg Pro Leu Glu Gly His Gly Ala Gly
 50 55 60

Val Gly Gly Arg Arg Arg Ser Gly Pro Leu Arg Thr Ser Ser Trp Gln
 65 70 75 80

Arg Ser Thr Lys Leu Pro Pro Pro Arg Arg Arg Ala Ser Ala Cys Gly
 85 90 95

Gly Leu Gly Leu Pro Arg Trp Pro Asp Lys Glu Val Leu Leu Glu Ala
 100 105 110

Glu Trp Arg Leu Val Arg Glu Met Arg Gly Glu Gly Leu Gly Arg Gln
 115 120 125

Pro His Glu Gly Ala Glu Arg Ser Arg Arg Gly Gln Leu Thr Val Phe
 130 135 140

Gln Leu Phe His Gln Leu Leu Leu Arg Gln Ala Thr Cys Arg Gly Leu
 145 150 155 160

Ala Glu Ala Val His Gly Gly Gly
 165

<210> 302
 <211> 32
 <212> PRT
 <213> Homo sapiens

<400> 302
 Pro Gly Val Pro Thr Gly Ile Arg Glu Ala Pro Glu Ser Pro Arg Thr
 1 5 10 15

Arg Ala Ser Gln Gly Ile Met Ala Ala Leu Phe Lys Lys Glu Val
 20 25 30

<210> 303
 <211> 28
 <212> PRT
 <213> Homo sapiens

<400> 303
 Phe Ile Leu Gly Gly Val Arg Gly Val Pro Arg Pro Leu Glu Gly His
 1 5 10 15

164

Gly Ala Gly Val Gly Gly Arg Arg Arg Ser Gly Pro
 20 25

<210> 304
 <211> 24
 <212> PRT
 <213> Homo sapiens

<400> 304
 Gly Leu Pro Arg Trp Pro Asp Lys Glu Val Leu Leu Glu Ala Glu Trp
 1 5 10 15

Arg Leu Val Arg Glu Met Arg Gly
 20

<210> 305
 <211> 23
 <212> PRT
 <213> Homo sapiens

<400> 305
 Gly Ala Glu Arg Ser Arg Arg Gly Gln Leu Thr Val Phe Gln Leu Phe
 1 5 10 15

His Gln Leu Leu Leu Arg Gln
 20

<210> 306
 <211> 15
 <212> PRT
 <213> Homo sapiens

<400> 306
 His Ala Ser Ala His Ala Ser Ala His Ala Ser Gly Cys Gly Ala
 1 5 10 15

<210> 307
 <211> 118
 <212> PRT
 <213> Homo sapiens

<400> 307
 Gln Gly Val Gly Val Ala Asp Glu Gly Gly Leu Glu Arg Gln Arg Val
 1 5 10 15

Asp Ala Gly Ala Arg Leu Gly His Met Gly Gln Pro Val Ala Phe Ser
 20 25 30

Thr Arg Gln Leu His Leu Ala Leu Pro Ala Pro Gly Thr Ala Gly Val
 35 40 45

Thr Val Pro His Pro His Ala Arg Glu Gly Val Val Gly Asp Leu Pro
 50 55 60

Leu Val Pro Asp Ala Glu Asp Pro Thr Val Gly Val Pro Ala Glu Gly
 65 70 75 80

Leu Leu Val Leu Gly His Val Val Glu Arg Ala Glu Leu Ile Leu Val
 85 90 95

Arg Gly Leu His Gln Ala Glu Ala Leu Ala Arg Glu Ser Glu Glu Met
 100 105 110

His Gly Ser Arg His Gly
 115

<210> 308

<211> 25

<212> PRT

<213> Homo sapiens

<400> 308

Glu Gly Gly Leu Glu Arg Gln Arg Val Asp Ala Gly Ala Arg Leu Gly
 1 5 10 15

His Met Gly Gln Pro Val Ala Phe Ser
 20 25

<210> 309

<211> 29

<212> PRT

<213> Homo sapiens

<400> 309

Leu Ala Leu Pro Ala Pro Gly Thr Ala Gly Val Thr Val Pro His Pro
 1 5 10 15

His Ala Arg Glu Gly Val Val Gly Asp Leu Pro Leu Val
 20 25

<210> 310

<211> 28

<212> PRT

<213> Homo sapiens

<400> 310

Pro Ala Glu Gly Leu Leu Val Leu Gly His Val Val Glu Arg Ala Glu
 1 5 10 15

Leu Ile Leu Val Arg Gly Leu His Gln Ala Glu Ala
 20 25

<210> 311

<211> 125

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (32)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 311

His Leu Phe Lys Phe Phe Tyr Thr Ile Ala Phe Met Gln Trp Phe Thr
 1 5 10 15

Glu Phe Met Glu Leu Phe Leu Ser Val Trp Glu Leu Ile Lys Thr Xaa
 20 25 30

Asn Leu Cys Phe Val Cys Phe Ser Glu His Lys Pro Gly Gln Leu Val
 35 40 45

Pro Ala Gly Pro Thr Ser Gln Leu Leu Cys Arg Ala Leu Gly Arg Val
 50 55 60

His Leu Cys Ser Pro Thr Thr Arg Ser Gln Thr Pro Thr Gln Ser Trp
 65 70 75 80

Val Thr Pro Gln Leu Leu Trp Arg Leu Gly Ser Gly Arg Leu Val Ala
 85 90 95

Gln Val Leu Gln Val Gly Ser Phe Cys Gly Pro Arg Val Gly Asp Ala
 100 105 110

Val Leu Gly Glu Gln Thr Phe Gln Pro Phe Asp Leu Leu
 115 120 125

<210> 312

<211> 29

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (23)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 312

Ala Phe Met Gln Trp Phe Thr Glu Phe Met Glu Leu Phe Leu Ser Val
 1 5 10 15

Trp Glu Leu Ile Lys Thr Xaa Asn Leu Cys Phe Val Cys
 20 25

<210> 313

<211> 26

<212> PRT

<213> Homo sapiens

<400> 313

Arg Ser Gln Thr Pro Thr Gln Ser Trp Val Thr Pro Gln Leu Leu Trp

WO 99/31117

PCT/US98/27059

1 5 10 15 167
Arg Leu Gly Ser Gly Arg Leu Val Ala Gln
20 25

<210> 314
<211> 39
<212> PRT
<213> Homo sapiens

<400> 314
Gly Ala Trp Gly Val Glu Val Val Ala Val Gly Ser Lys Ala Gly Cys
1 5 10 15
Leu Val Tyr Gln Leu Cys Asp Leu Lys Gln Ile Thr Phe Phe Phe Arg
20 25 30
Ala Ser Val Cys Leu Ser Val
35

<210> 315
<211> 194
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (61)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (95)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (116)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (129)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (131)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (132)
<223> Xaa equals any of the naturally occurring L-amino acids

SUBSTITUTE SHEET (RULE 26)

<220>

<221> SITE

<222> (163)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (187)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 315

Pro Ala Ser Leu Gly Ser Ser Trp Gly Gln Lys Leu Arg Gly Gly Thr
 1 5 10 15

Arg Lys Ser Phe Gln Glu Leu Ser Pro Ser Ser Ala Pro Pro Ala Cys
 20 25 30

Leu Pro Gln Pro Pro Ala Ser Thr Trp Leu Ser Ser Trp Pro Arg Pro
 35 40 45

Pro Cys Trp Pro Pro Met Cys Ser Trp Ala Leu Gly Xaa Cys Phe Cys
 50 55 60

Pro Ala Thr Gly Gln Trp Leu Pro Thr Ser Cys Cys Leu Trp Trp Cys
 65 70 75 80

Pro Asp Ala Gly Gly Arg Gln Lys His Phe Arg Ser Arg Trp Xaa Thr
 85 90 95

Ser Trp Glu Thr Trp Gln Pro Tyr Leu Thr Gly Leu Ile Ser Ser Val
 100 105 110

Leu Arg Ala Xaa Arg Pro Asp Ser Tyr Leu Gln Arg Phe Arg Ser Leu
 115 120 125

Xaa Gln Xaa Xaa Leu Cys Cys Ala Phe Val Ile Ala Leu Gly Gly Gly
 130 135 140

Cys Phe Leu Leu Thr Ala Leu Tyr Leu Glu Arg Asp Glu Thr Arg Ala
 145 150 155 160

Trp Gln Xaa Val Thr Gly Thr Pro Asp Ser Asn Asp Val Asp Ser Asn
 165 170 175

Asp Leu Glu Arg Gln Gly Leu Leu Ser Gly Xaa Gly Ala Ser Thr Glu
 180 185 190

Glu Pro

<210> 316

<211> 26

<212> PRT

<213> Homo sapiens

<400> 316

169

Leu Arg Gly Gly Thr Arg Lys Ser Phe Gln Glu Leu Ser Pro Ser Ser
 1 5 10 15

Ala Pro Pro Ala Cys Leu Pro Gln Pro Pro
 20 25

<210> 317

<211> 28

<212> PRT

<213> Homo sapiens

<400> 317

Ala Thr Gly Gln Trp Leu Pro Thr Ser Cys Cys Leu Trp Trp Cys Pro
 1 5 10 15

Asp Ala Gly Gly Arg Gln Lys His Phe Arg Ser Arg
 20 25

<210> 318

<211> 22

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (21)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 318

Gly Gly Cys Phe Leu Leu Thr Ala Leu Tyr Leu Glu Arg Asp Glu Thr
 1 5 10 15

Arg Ala Trp Gln Xaa Val
 20

<210> 319

<211> 124

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (38)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (72)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (76)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>
 <221> SITE
 <222> (93)
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>
 <221> SITE
 <222> (105)
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>
 <221> SITE
 <222> (106)
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>
 <221> SITE
 <222> (107)
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>
 <221> SITE
 <222> (108)
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>
 <221> SITE
 <222> (109)
 <223> Xaa equals any of the naturally occurring L-amino acids

<400> 319
 Ala Pro His Leu Arg Leu Gln Pro Ala Cys His Ser Pro Leu Pro Leu
 1 5 10 15
 Pro Gly Ser Arg Pro Gly Pro Asp His Pro Ala Gly Leu Leu Cys Val
 20 25 30
 Pro Gly Pro Trp Gly Xaa Ala Ser Val Leu Gln Leu Gly Ser Gly Cys
 35 40 45
 Arg His Pro Ala Val Cys Gly Gly Ala Gln Met Pro Gly Asp Gly Arg
 50 55 60
 Ser Thr Ser Asp His Gly Gly Xaa His Pro Gly Xaa Pro Gly Ser Pro
 65 70 75 80
 Ile Ser Gln Asp Leu Ser Leu Val Ser Cys Gly Pro Xaa Ala Leu Thr
 85 90 95
 Pro Ile Cys Ser Ala Ser Ala Ala Xaa Xaa Xaa Xaa Xaa Cys Ala Ala
 100 105 110
 Pro Leu Ser Ser Pro Trp Gly Ala Ala Ala Ser Cys
 115 120

171

<210> 320
 <211> 25
 <212> PRT
 <213> Homo sapiens

<400> 320
 Pro Ala Cys His Ser Pro Leu Pro Leu Pro Gly Ser Arg Pro Gly Pro
 1 5 10 15
 Asp His Pro Ala Gly Leu Leu Cys Val
 20 25

<210> 321
 <211> 26
 <212> PRT
 <213> Homo sapiens

<400> 321
 Ser Gly Cys Arg His Pro Ala Val Cys Gly Gly Ala Gln Met Pro Gly
 1 5 10 15
 Asp Gly Arg Ser Thr Ser Asp His Gly Gly
 20 25

<210> 322
 <211> 95
 <212> PRT
 <213> Homo sapiens

<400> 322
 Gly Leu Lys Val Met Glu Ile Cys Ser Leu Thr Phe Leu Glu Ala Thr
 1 5 10 15
 Asn Leu Gln Ser Arg Cys Gln Gln Ala Met Leu Pro Leu Lys Ala Leu
 20 25 30
 Arg Lys Asn Pro Phe Leu Leu Leu Pro Ser Phe Asp Gly Cys Cys Gln
 35 40 45
 Ser Leu Ala Phe Pro Gly Leu Trp Leu Gln His Ser Asn Leu Cys Leu
 50 55 60
 Asn His His Met Thr Phe Leu Val Tyr Leu Leu Cys Val Ser Val Phe
 65 70 75 80
 Lys Tyr Phe Phe Pro Phe Ser Cys Thr Tyr Thr Ser His Trp Ile
 85 90 95

<210> 323
 <211> 22
 <212> PRT
 <213> Homo sapiens

<400> 323

172

Ile Cys Ser Leu Thr Phe Leu Glu Ala Thr Asn Leu Gln Ser Arg Cys
 1 5 10 15

Gln Gln Ala Met Leu Pro
 20

<210> 324

<211> 26

<212> PRT

<213> Homo sapiens

<400> 324

Gly Leu Trp Leu Gln His Ser Asn Leu Cys Leu Asn His His Met Thr
 1 5 10 15

Phe Leu Val Tyr Leu Leu Cys Val Ser Val
 20 25

<210> 325

<211> 37

<212> PRT

<213> Homo sapiens

<400> 325

Pro Phe Pro Leu Leu Pro Pro Lys Arg Arg Gly Leu Leu Tyr His Leu
 1 5 10 15

Ile Gln Lys Ser Thr Leu Gly Leu Val Val Trp Phe Arg Glu His Leu
 20 25 30

Asp Ser Arg Ser Gln
 35

<210> 326

<211> 78

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (3)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (46)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (48)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (65)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 326

Arg Gly Xaa Pro Ser Trp Pro Met His Thr Leu Val Tyr Ala Gln His
 1 5 10 15

Ser Thr Thr His Thr Pro Leu Ile Gln Pro Gln Trp Thr Gln Val Ile
 20 25 30

Asp Gln Pro Pro Gly Ile Thr His Gln Phe Cys Val Arg Xaa Cys Xaa
 35 40 45

Cys Pro Thr Leu Glu Ser Cys Val Gln Glu Cys Val Thr Arg Ser Arg
 50 55 60

Xaa Lys Pro Thr Thr Gly Val Pro Gly Pro Gln Arg Leu Ala
 65 70 75

<210> 327

<211> 24

<212> PRT

<213> Homo sapiens

<400> 327

Thr Pro Leu Ile Gln Pro Gln Trp Thr Gln Val Ile Asp Gln Pro Pro
 1 5 10 15

Gly Ile Thr His Gln Phe Cys Val
 20

<210> 328

<211> 104

<212> PRT

<213> Homo sapiens

<400> 328

Ala Leu Gly Pro Ser Gln Thr Cys Asp Leu Asp Val Trp Leu Val Ala
 1 5 10 15

Lys Pro Ser Phe Phe Arg Gly Pro Gln Gly Ile His Tyr Phe Ser Leu
 20 25 30

Trp Arg Arg Lys Pro Leu Ser His Trp Val Ser Ile Trp Gln Leu Gln
 35 40 45

Gly Gln Glu Thr Met Pro Ala Met Leu Arg Ser Arg Pro Ala Gly Gln
 50 55 60

Ala Thr Val Ala Thr Gly Pro Pro Arg Gly Ser Pro Ser Pro Gln Asp
 65 70 75 80

Leu Pro Ser Tyr His Arg Lys Gln Val Glu Ser Ser His Arg His Ser
 85 90 95

Trp Glu Pro Ala Ser Gln Ser Gln
100

<210> 329
<211> 28
<212> PRT
<213> Homo sapiens

<400> 329
Cys Asp Leu Asp Val Trp Leu Val Ala Lys Pro Ser Phe Phe Arg Gly
1 5 10 15

Pro Gln Gly Ile His Tyr Phe Ser Leu Trp Arg Arg
20 25

<210> 330
<211> 28
<212> PRT
<213> Homo sapiens

<400> 330
Ala Gly Gln Ala Thr Val Ala Thr Gly Pro Pro Arg Gly Ser Pro Ser
1 5 10 15

Pro Gln Asp Leu Pro Ser Tyr His Arg Lys Gln Val
20 25

<210> 331
<211> 79
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (1)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (5)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (15)
<223> Xaa equals any of the naturally occurring L-amino acids

<400> 331
Xaa Gly Asp Thr Xaa Thr Gln Asn Ser Arg His Asp Thr Pro Xaa Leu
1 5 10 15

Ile Asp Tyr Tyr Arg Glu Ser Cys Thr Leu Gln Tyr Arg Pro Glu Phe
20 25 30

Pro Gly Arg Pro Thr Arg Pro Arg Gly Ser Cys Pro Gln Tyr Pro Gly
 35 40 45

Pro Ala Ile Pro Arg Thr Ser Trp Ala Leu Gly Glu Gly Asp Ala Ala
 50 55 60

Pro Arg Gly Ala His His Pro Arg Arg Ala Asp Val Pro Leu Gly
 65 70 75

<210> 332

<211> 30

<212> PRT

<213> Homo sapiens

<400> 332

Tyr Arg Glu Ser Cys Thr Leu Gln Tyr Arg Pro Glu Phe Pro Gly Arg
 1 5 10 15

Pro Thr Arg Pro Arg Gly Ser Cys Pro Gln Tyr Pro Gly Pro
 20 25 30

<210> 333

<211> 155

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (72)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 333

Gly Lys Leu Tyr Ala Ala Val Pro Ser Gly Ile Pro Gly Ser Thr His
 1 5 10 15

Ala Ser Ala Arg Leu Met Pro Pro Val Ser Arg Ser Ser Tyr Ser Glu
 20 25 30

Asp Ile Val Gly Ser Arg Arg Arg Arg Ser Ser Ser Gly Ser Pro
 35 40 45

Pro Ser Pro Gln Ser Arg Cys Ser Ser Trp Asp Gly Cys Ser Arg Ser
 50 55 60

His Ser Arg Gly Arg Glu Gly Xaa Arg Pro Pro Trp Ser Glu Leu Asp
 65 70 75 80

Val Gly Ala Leu Tyr Pro Phe Ser Arg Ser Gly Ser Arg Gly Arg Leu
 85 90 95

Pro Arg Phe Arg Asn Tyr Ala Phe Ala Ser Ser Trp Ser Thr Ser Tyr
 100 105 110

Ser Gly Tyr Arg Tyr His Arg Ala Leu Leu Cys Arg Arg Thr Ala Val

176

115

120

125

Ser Gly Arg Leu Arg Glu Gly Arg Glu Pro Ser Ala Glu Glu Ala Glu
 130 135 140

Gly Glu Arg Glu Asp Trp Gly Ile Gly Ser Ala
 145 150 155

<210> 334
 <211> 23
 <212> PRT
 <213> Homo sapiens

<400> 334
 Ser Gly Ile Pro Gly Ser Thr His Ala Ser Ala Arg Leu Met Pro Pro
 1 5 10 15

Val Ser Arg Ser Ser Tyr Ser
 20

<210> 335
 <211> 29
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (13)
 <223> Xaa equals any of the naturally occurring L-amino acids

<400> 335
 Gly Cys Ser Arg Ser His Ser Arg Gly Arg Glu Gly Xaa Arg Pro Pro
 1 5 10 15

Trp Ser Glu Leu Asp Val Gly Ala Leu Tyr Pro Phe Ser
 20 25

<210> 336
 <211> 25
 <212> PRT
 <213> Homo sapiens

<400> 336
 Thr Ala Val Ser Gly Arg Leu Arg Glu Gly Arg Glu Pro Ser Ala Glu
 1 5 10 15

Glu Ala Glu Gly Glu Arg Glu Asp Trp
 20 25

<210> 337
 <211> 134
 <212> PRT
 <213> Homo sapiens

<220>

<221> SITE

<222> (17)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (77)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 337

Arg	Ile	Arg	Lys	Ala	Ala	Val	Gln	Ile	Pro	Thr	Arg	Lys	Asn	Ile	Gly
1				5					10					15	

Xaa	Arg	Arg	Pro	Val	Val	Gln	Glu	Thr	Arg	Lys	Lys	Glu	Arg	Ile	Ser
			20					25					30		

Arg	Leu	Lys	Glu	Ser	Ile	Gly	Asn	Ile	Leu	Ile	Val	Thr	Val	Thr	Gln
		35					40					45			

Ser	Leu	Thr	Gln	Ile	Leu	Thr	Leu	Met	Met	Ile	Lys	Arg	Glu	Leu	Lys
	50					55					60				

Pro	Arg	Arg	Lys	Arg	Arg	Lys	Arg	Asn	Thr	Lys	Gln	Xaa	Lys	Arg	Arg
	65				70					75					80

Ile	Arg	Lys	Pro	Lys	Lys	Asn	Pro	Val	Thr	Gln	Ala	Val	Lys	Thr	Gln
				85					90					95	

Lys	Arg	Thr	Cys	Gln	Lys	Leu	Pro	Gly	Met	Glu	Gln	Pro	Asn	Val	Ala
		100						105					110		

Asp	Thr	Met	Asp	Leu	Ile	Gly	Pro	Glu	Ala	Pro	Ile	Asn	Thr	Tyr	Leu
		115					120					125			

Phe	Lys	Met	Lys	Asn	Leu
		130			

<210> 338

<211> 28

<212> PRT

<213> Homo sapiens

<400> 338

Thr	Arg	Lys	Lys	Glu	Arg	Ile	Ser	Arg	Leu	Lys	Glu	Ser	Ile	Gly	Asn
1				5					10					15	

Ile	Leu	Ile	Val	Thr	Val	Thr	Gln	Ser	Leu	Thr	Gln
			20				25				

<210> 339

<211> 28

<212> PRT

<213> Homo sapiens

<400> 339

Val Lys Thr Gln Lys Arg Thr Cys Gln Lys Leu Pro Gly Met Glu Gln
 1 5 10 15

Pro Asn Val Ala Asp Thr Met Asp Leu Ile Gly Pro
 20 25

<210> 340

<211> 80

<212> PRT

<213> Homo sapiens

<400> 340

Leu Pro Phe Thr Leu Lys Pro Lys Met Val Lys Ile Pro Phe Ser Ser
 1 5 10 15

Arg Leu Ile Asn Asn Leu Gln Tyr Ile Asp Cys Ile Leu Ser Leu
 20 25 30

Lys Arg Cys Glu Glu Ile Leu Leu Met Trp His Gly Leu Leu Leu Cys
 35 40 45

Leu Ala Ser Val Phe Leu Glu Leu Arg Gly Asp Arg Pro Pro Leu Leu
 50 55 60

Ala Ser Leu Leu Glu Pro His Lys Met Pro Leu His Ser Ser Ser Leu
 65 70 75 80

<210> 341

<211> 24

<212> PRT

<213> Homo sapiens

<400> 341

Leu Lys Pro Lys Met Val Lys Ile Pro Phe Ser Ser Arg Leu Ile Asn
 1 5 10 15

Asn Asn Leu Gln Tyr Ile Asp Cys
 20

<210> 342

<211> 23

<212> PRT

<213> Homo sapiens

<400> 342

Ser Leu Lys Arg Cys Glu Glu Ile Leu Leu Met Trp His Gly Leu Leu
 1 5 10 15

Leu Cys Leu Ala Ser Val Phe

20

<210> 343
<211> 21
<212> PRT
<213> Homo sapiens

<400> 343
Leu Arg Gly Asp Arg Pro Pro Leu Leu Ala Ser Leu Leu Glu Pro His
1 5 10 15

Lys Met Pro Leu His
20

<210> 344
<211> 79
<212> PRT
<213> Homo sapiens

<400> 344
Leu Gln Met His Thr Gly Ser Gly Phe Lys Gly Lys Ser Cys Glu Val
1 5 10 15

Ala Phe Tyr Val Ala Gln Ala Glu Lys Pro Gly Glu Gly Ala Tyr Leu
20 25 30

His Gly Ala Gln Glu Thr Gln Lys Gln Gly Ile Glu Ala Asp His Ala
35 40 45

Thr Leu Arg Gly Ser Pro His Ser Val Ser Lys Thr Lys Tyr Asn Leu
50 55 60

Tyr Ile Ala Asn Tyr Tyr Leu Leu Ala Trp Arg Lys Met Glu Ser
65 70 75

<210> 345
<211> 20
<212> PRT
<213> Homo sapiens

<400> 345
Cys Glu Val Ala Phe Tyr Val Ala Gln Ala Glu Lys Pro Gly Glu Gly
1 5 10 15

Ala Tyr Leu His
20

<210> 346
<211> 23
<212> PRT
<213> Homo sapiens

<400> 346

WO 99/31117

PCT/US98/27059

180

Ala Thr Leu Arg Gly Ser Pro His Ser Val Ser Lys Thr Lys Tyr Asn
1 5 10 15

Leu Tyr Ile Ala Asn Tyr Tyr
20

<210> 347

<211> 65

<212> PRT

<213> Homo sapiens

<400> 347

Leu Ser Ala Ser Leu Leu Asp Arg Tyr Pro Ala Ser Glu Ser Asn Asn
1 5 10 15

Tyr Ile Phe Asn Phe Val Leu Tyr Met Leu His Phe Leu Ala Gly Thr
20 25 30

Leu Phe Ser Leu Phe Pro Asp Phe Glu Leu Ser Pro Arg Ser Ala Thr
35 40 45

Leu Phe Pro Asp Leu Arg Thr Val Gln Leu Leu Ser Ser Arg Pro His
50 55 60

Leu

65

<210> 348

<211> 23

<212> PRT

<213> Homo sapiens

<400> 348

Leu Leu Asp Arg Tyr Pro Ala Ser Glu Ser Asn Asn Tyr Ile Phe Asn
1 5 10 15

Phe Val Leu Tyr Met Leu His
20

<210> 349

<211> 20

<212> PRT

<213> Homo sapiens

<400> 349

Phe Pro Asp Phe Glu Leu Ser Pro Arg Ser Ala Thr Leu Phe Pro Asp
1 5 10 15

Leu Arg Thr Val
20

<210> 350

<211> 85

SUBSTITUTE SHEET (RULE 26)

<212> PRT

<213> Homo sapiens

<400> 350

Asn Gly Gly Phe Tyr Asp Val Ser Phe Lys Gln Ala Gly Leu Ile Glu
 1 5 10 15

Phe Leu Cys Ile Ile Tyr Phe Tyr Pro Met Ala His Val Ile Cys Gly
 20 25 30

Ser Arg Phe Thr Ile Val Arg Thr Ile Pro Val His Tyr Val Gly Glu
 35 40 45

Tyr Phe Ile Lys Ser Ser Ile Trp Ile Leu Tyr Arg Ile Asn Glu Arg
 50 55 60

Thr Ala Thr Lys Lys Ala Ala Ser Asp Phe Gln Lys Asn Phe Arg Cys
 65 70 75 80

Phe Leu Asp Ala Phe
 85

<210> 351

<211> 19

<212> PRT

<213> Homo sapiens

<400> 351

Lys Gln Ala Gly Leu Ile Glu Phe Leu Cys Ile Ile Tyr Phe Tyr Pro
 1 5 10 15

Met Ala His

<210> 352

<211> 23

<212> PRT

<213> Homo sapiens

<400> 352

Tyr Phe Ile Lys Ser Ser Ile Trp Ile Leu Tyr Arg Ile Asn Glu Arg
 1 5 10 15

Thr Ala Thr Lys Lys Ala Ala
 20

<210> 353

<211> 22

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (4)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (7)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (9)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 353

Ser	Pro	Arg	Xaa	Gly	Arg	Xaa	Phe	Xaa	Thr	Ser	Arg	Lys	Gln	Ile	Ser
1				5					10					15	

Gly	Phe	Leu	Glu	Phe	Asp
					20

<210> 354

<211> 56

<212> PRT

<213> Homo sapiens

<400> 354

Met	Lys	His	Ala	Ala	Phe	Gly	Leu	Ile	Pro	Leu	Val	Lys	Glu	Ile	Tyr
1				5					10					15	

Arg	Tyr	Leu	Lys	Ile	Lys	Ser	Lys	Leu	Leu	Ile	Gly	Ser	Gly	Lys	Cys
		20						25					30		

Gln	Leu	Gln	Pro	Glu	Trp	Leu	Gln	Thr	Ser	Leu	Ile	Asn	Ser	Ser	Leu
		35					40					45			

Leu	Met	Asp	Trp	Leu	Thr	Pro	Tyr
	50					55	

<210> 355

<211> 29

<212> PRT

<213> Homo sapiens

<400> 355

Ile	Tyr	Arg	Tyr	Leu	Lys	Ile	Lys	Ser	Lys	Leu	Leu	Ile	Gly	Ser	Gly
1				5					10					15	

Lys	Cys	Gln	Leu	Gln	Pro	Glu	Trp	Leu	Gln	Thr	Ser	Leu
			20					25				

<210> 356

<211> 68

<212> PRT

<213> Homo sapiens

183

<400> 356

Gln Leu Gly Leu Pro Trp Asp Gln Ser Lys Gly Pro Arg Lys Asn Gly
 1 5 10 15

Leu Ser Met Cys Gly Ser Val Tyr Ser Thr Ile Trp Ser Leu Ile Ala
 20 25 30

Ser Arg Arg Glu Glu Thr Ile Arg Val Ile Val Leu Tyr Ile Gln Ser
 35 40 45

Pro Asn Ile Asn Thr Arg His Ile Ser Lys Arg Gly Leu Asn Lys Ala
 50 55 60

Leu Thr Asn Pro
 65

<210> 357

<211> 21

<212> PRT

<213> Homo sapiens

<400> 357

Ser Lys Gly Pro Arg Lys Asn Gly Leu Ser Met Cys Gly Ser Val Tyr
 1 5 10 15

Ser Thr Ile Trp Ser
 20

<210> 358

<211> 17

<212> PRT

<213> Homo sapiens

<400> 358

Gln Ser Pro Asn Ile Asn Thr Arg His Ile Ser Lys Arg Gly Leu Asn
 1 5 10 15

Lys

<210> 359

<211> 19

<212> PRT

<213> Homo sapiens

<400> 359

His Pro Gln Thr Ser Ala Gly Gly Phe Pro Leu His Gln Gly Leu Pro
 1 5 10 15

Thr Val Ser

<210> 360

<211> 117
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (110)
 <223> Xaa equals any of the naturally occurring L-amino acids

<400> 360

Pro	Ser	Trp	Phe	Pro	Glu	Leu	Ser	Pro	Trp	Pro	Leu	Lys	Thr	Leu	Lys
1				5					10					15	
Lys	Arg	Arg	Gln	Met	Arg	Leu	Arg	Arg	Arg	Gly	Arg	Leu	Cys	Arg	Leu
			20					25					30		
Ser	Pro	Ala	Thr	Thr	Thr	Thr	Ala	Asp	Thr	Cys	Arg	Cys	Pro	Ala	Arg
		35					40					45			
Ser	Tyr	Arg	Gly	Ser	Gly	Arg	Arg	Pro	Ala	Cys	Ala	Gln	Asp	Ser	Pro
	50				55						60				
Ala	Pro	Pro	Ser	Arg	Pro	Thr	Arg	Arg	Ala	Trp	Glu	Lys	Cys	Ala	Leu
65					70					75					80
Arg	Pro	Lys	Arg	Ala	Ala	Gln	Trp	Ser	Thr	Gly	Val	Pro	Pro	Ser	Pro
				85					90					95	
Arg	Ser	Ser	Thr	Thr	Gly	Cys	Cys	Phe	Gly	Thr	Ala	Ala	Xaa	Cys	Ala
			100					105						110	
Glu	Gly	Ala	Arg	Arg											
			115												

<210> 361
 <211> 22
 <212> PRT
 <213> Homo sapiens

<400> 361

Thr	Thr	Thr	Ala	Asp	Thr	Cys	Arg	Cys	Pro	Ala	Arg	Ser	Tyr	Arg	Gly
1				5						10				15	
Ser	Gly	Arg	Arg	Pro	Ala										
			20												

<210> 362
 <211> 24
 <212> PRT
 <213> Homo sapiens

<400> 362

Pro	Ser	Arg	Pro	Thr	Arg	Arg	Ala	Trp	Glu	Lys	Cys	Ala	Leu	Arg	Pro
1				5					10					15	

Lys Arg Ala Ala Gln Trp Ser Thr
20

<210> 363
<211> 20
<212> PRT
<213> Homo sapiens

<400> 363
Ala Arg Gly Val Leu Asn Leu Arg Asn Arg Phe Glu Cys Phe Ser Ile
1 5 10 15

Ile Glu Thr Val
20

<210> 364
<211> 69
<212> PRT
<213> Homo sapiens

<400> 364
Ile Gly Gln Leu Val Met Lys Ser Ile Cys His Phe Gln Arg Leu Leu
1 5 10 15

Ser Val Ala Ile Asp Phe Ala Ser Gln Phe Leu Lys Asn Tyr Ile Phe
20 25 30

Ser Ser Thr His Ser Ser Lys Ala Gly Phe Ser Val Val Cys Ser Leu
35 40 45

Pro Lys Trp Leu Tyr Thr Asp Gly Met Glu Met Val Leu Lys Ile Thr
50 55 60

His Lys Leu Ser Phe
65

<210> 365
<211> 24
<212> PRT
<213> Homo sapiens

<400> 365
Gln Arg Leu Leu Ser Val Ala Ile Asp Phe Ala Ser Gln Phe Leu Lys
1 5 10 15

Asn Tyr Ile Phe Ser Ser Thr His
20

<210> 366
<211> 12
<212> PRT
<213> Homo sapiens

186

<400> 366

Leu Met Lys Thr Ala Ser Arg Met Leu Leu Leu Glu
 1 5 10

<210> 367

<211> 25

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (3)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (6)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 367

Ala Thr Xaa Trp Asp Xaa Pro Gly Cys Arg Asn Ser Ala Arg Gly Glu
 1 5 10 15

Arg Leu His Val Gly Asp Ala Pro Trp
 20 25

<210> 368

<211> 109

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (102)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (105)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 368

Ala Arg Asp Glu Arg Arg Glu Val Leu Lys Thr Leu Met Arg Leu Ser
 1 5 10 15

Thr Gln Arg Pro Gln Ala Phe Leu Pro Ser Gln Ser Trp Phe Val Arg
 20 25 30

Leu Gln Lys Ala Gly Glu Gly Ala Leu Lys Gln Glu Asn Ser Leu Thr
 35 40 45

Ile Gln Asn Cys Leu Leu Cys Leu Pro Arg Val His Arg Gln Arg Pro
 50 55 60

Thr Pro Pro Gln Pro Gln Arg Gly Asn Thr Glu Ala Ser Val Leu Gln

PCT/US98/27059

```
<210> 369
<211> 22
<212> PRT
<213> Homo sapiens
```

```
<210> 370
<211> 25
<212> PRT
<213> Homo sapiens
```

```
<210> 371
<211> 21
<212> PRT
<213> Homo sapiens
```

```
<210> 372
<211> 9
<212> PRT
<213> Homo sapiens
```

SUBSTITUTE SHEET (RULE 26)

<210> 373
 <211> 63
 <212> PRT
 <213> Homo sapiens

<400> 373
 Pro Tyr Ile Asn Thr Gln Met Cys Val Ser Ser Arg Asn Lys Phe Cys
 1 5 10 15
 Ile Ser Gly His Gln Lys Tyr Asp Ser His Gly Arg Glu Thr Arg Phe
 20 25 30
 Glu Met His Lys Ala Arg Ala Ser Ser Trp Lys Asn Ile Leu Lys Ile
 35 40 45
 Arg Ser Leu Lys Ile Ile Ser Arg Gly Phe Glu Ile Thr Asn Ala
 50 55 60

<210> 374
 <211> 27
 <212> PRT
 <213> Homo sapiens

<400> 374
 Lys Phe Cys Ile Ser Gly His Gln Lys Tyr Asp Ser His Gly Arg Glu
 1 5 10 15
 Thr Arg Phe Glu Met His Lys Ala Arg Ala Ser
 20 25

<210> 375
 <211> 84
 <212> PRT
 <213> Homo sapiens

<400> 375
 His Thr Leu Leu Glu Ile Ala Asn Pro Leu Gln Ala Ala Val Leu Gly
 1 5 10 15
 Ala Ser Ser Ile His Pro Ser Ile His Thr Ser Thr His Leu Met Phe
 20 25 30
 Met Gly Leu Lys Trp Thr Glu Leu His His Ser Pro Asp Ser Val Gln
 35 40 45
 Gly Ala Gly Ala Ala Glu Ala Ala Gln Thr Arg His Ser Leu Arg Pro
 50 55 60
 Gly Arg Gly Arg Glu Arg His Asp Cys Thr Leu Lys Asn Leu Thr Leu
 65 70 75 80
 Phe Ile Ile Cys

<210> 376
 <211> 22
 <212> PRT
 <213> Homo sapiens

<400> 376
 Asn Pro Leu Gln Ala Ala Val Leu Gly Ala Ser Ser Ile His Pro Ser
 1 5 10 15

Ile His Thr Ser Thr His
 20

<210> 377
 <211> 17
 <212> PRT
 <213> Homo sapiens

<400> 377
 Ser Leu Arg Pro Gly Arg Gly Arg Glu Arg His Asp Cys Thr Leu Lys
 1 5 10 15

Asn

<210> 378
 <211> 52
 <212> PRT
 <213> Homo sapiens

<400> 378
 Ala Glu Asn Val His Cys Thr Pro Ala Trp Glu Thr Gly Arg Asp Ser
 1 5 10 15

Glu Asp Gly Lys Gly Arg Glu Gly Met Gly Arg Asp Arg Lys Gly Trp
 20 25 30

Asp Gly Thr Gly Leu Asp Gly Thr Gly Trp Glu Gly Lys Arg Glu Arg
 35 40 45

Asn Val Pro Ala
 50

<210> 379
 <211> 26
 <212> PRT
 <213> Homo sapiens

<400> 379
 Gly Arg Asp Ser Glu Asp Gly Lys Gly Arg Glu Gly Met Gly Arg Asp
 1 5 10 15

Arg Lys Gly Trp Asp Gly Thr Gly Leu Asp
 20 25

<210> 380
 <211> 14
 <212> PRT
 <213> Homo sapiens

<400> 380
 Thr Ser Leu Gly Asp Leu Trp Asp Tyr Asn Asn Ser Ser His
 1 5 10

<210> 381
 <211> 66
 <212> PRT
 <213> Homo sapiens

<400> 381
 Asp Arg Arg Ile Ile Arg Thr Arg Glu Ala Ala Val Ala Val Ser Arg
 1 5 10 15
 Glu Arg Pro Leu His Ser Ser Leu Gly Asn Arg Glu Arg Leu Arg Arg
 20 25 30
 Trp Glu Gly Thr Gly Arg Asp Gly Lys Gly Gln Glu Gly Met Gly Arg
 35 40 45
 Asp Gly Thr Gly Trp Asp Gly Met Gly Arg Glu Glu Arg Lys Lys Cys
 50 55 60
 Pro Ser
 65

<210> 382
 <211> 25
 <212> PRT
 <213> Homo sapiens

<400> 382
 Arg Pro Leu His Ser Ser Leu Gly Asn Arg Glu Arg Leu Arg Arg Trp
 1 5 10 15
 Glu Gly Thr Gly Arg Asp Gly Lys Gly
 20 25

<210> 383
 <211> 9
 <212> PRT
 <213> Homo sapiens

<400> 383
 Asn Gln Ser Trp Gly Pro Met Gly Leu
 1 5

<210> 384
<211> 59
<212> PRT
<213> Homo sapiens

<400> 384
Gly Gly Gly Gly Cys Ser Glu Pro Arg Thr Ser Ile Ala Leu Gln Pro
1 5 10 15
Gly Lys Gln Gly Glu Thr Pro Lys Met Gly Arg Asp Gly Lys Gly Trp
20 25 30
Glu Gly Thr Gly Arg Asp Gly Thr Gly Arg Asp Trp Met Gly Arg Asp
35 40 45
Gly Lys Gly Arg Glu Lys Glu Met Ser Gln Gln
50 55

<210> 385
<211> 24
<212> PRT
<213> Homo sapiens

<400> 385
Lys Gln Gly Glu Thr Pro Lys Met Gly Arg Asp Gly Lys Gly Trp Glu
1 5 10 15
Gly Thr Gly Arg Asp Gly Thr Gly
20

<210> 386
<211> 32
<212> PRT
<213> Homo sapiens

<400> 386
Pro Val Leu Gly Thr Tyr Gly Thr Ile Thr Thr Pro Val Thr Glu Leu
1 5 10 15
Thr Lys Gly Gln Glu Lys Glu Gly Gly Val Glu Thr Val Leu Tyr Glu
20 25 30

<210> 387
<211> 11
<212> PRT
<213> Homo sapiens

<400> 387
Lys Ile Val Phe Ile Asp Gln Lys Trp Ser Lys
1 5 10

<210> 388
 <211> 70
 <212> PRT
 <213> Homo sapiens

<400> 388
 Cys Ser Leu Phe Trp Gly Ile Leu Phe Leu Ser Arg Leu Arg Ile His
 1 5 10 15
 Leu Phe Leu Ser Leu Lys Pro Cys Met Cys Leu Arg Pro Ile Asp Ile
 20 25 30
 Leu Ser His Phe Leu Asp Ile Phe Val Thr Ser Val Leu Ser Glu Leu
 35 40 45
 Glu Lys Ser Ser Leu Lys Thr Thr Glu Thr Phe Ser Phe Ala Val Phe
 50 55 60
 Leu Leu Leu Met Met Asn
 65 70

<210> 389
 <211> 26
 <212> PRT
 <213> Homo sapiens

<400> 389
 Leu Ser Arg Leu Arg Ile His Leu Phe Leu Ser Leu Lys Pro Cys Met
 1 5 10 15
 Cys Leu Arg Pro Ile Asp Ile Leu Ser His
 20 25

<210> 390
 <211> 22
 <212> PRT
 <213> Homo sapiens

<400> 390
 Val Leu Ser Glu Leu Glu Lys Ser Ser Leu Lys Thr Thr Glu Thr Phe
 1 5 10 15
 Ser Phe Ala Val Phe Leu
 20

<210> 391
 <211> 8
 <212> PRT
 <213> Homo sapiens

<400> 391
 Thr Leu Phe Arg Tyr Ile Leu His
 1 5

<210> 392
<211> 14
<212> PRT
<213> Homo sapiens

<400> 392
Gly Thr Ser Phe Ser Val Leu Ser Leu Ile His Asp Thr Gly
1 5 10

<210> 393
<211> 63
<212> PRT
<213> Homo sapiens

<400> 393
Val Leu Ile Ser Ala Ser Thr Ile Gly Ser Arg Thr Ser Gly Ala Gln
1 5 10 15
Gly Met Glu Lys Met Thr Ile Pro Thr Leu Ala Val Gly Glu Pro Lys
20 25 30
Thr Pro Glu Lys Ser Lys Cys Ser Leu Lys Gln Cys Phe Ser Ser Cys
35 40 45
Asn Val His Ile Asp His Leu Gly Leu Leu Leu Lys Cys Lys Phe
50 55 60

<210> 394
<211> 23
<212> PRT
<213> Homo sapiens

<400> 394
Ala Ser Thr Ile Gly Ser Arg Thr Ser Gly Ala Gln Gly Met Glu Lys
1 5 10 15
Met Thr Ile Pro Thr Leu Ala
20

<210> 395
<211> 27
<212> PRT
<213> Homo sapiens

<400> 395
Gly Glu Pro Lys Thr Pro Glu Lys Ser Lys Cys Ser Leu Lys Gln Cys
1 5 10 15
Phe Ser Ser Cys Asn Val His Ile Asp His Leu
20 25

<210> 396
 <211> 101
 <212> PRT
 <213> Homo sapiens

<400> 396
 Arg Ile Arg Ser Gln Asp Leu Ala Ile Met Thr Thr Cys Phe Lys Lys
 1 5 10 15
 Tyr Glu Phe Ser Phe Phe Val Leu Gly Phe Leu Arg Arg Trp Gly Ala
 20 25 30
 Thr Leu Cys Leu Gly Phe Thr Ser Phe Ala Ile Lys Phe His Pro Ser
 35 40 45
 Ser Leu Cys Ser Glu Lys Glu Gly Lys Asp Phe Ser Gly Phe Ala Leu
 50 55 60
 Ser Ile His Gly Pro Glu Arg Lys Lys Glu Glu Gly Trp Ala Arg Trp
 65 70 75 80
 Leu Thr Pro Val Val Pro Val Leu Trp Glu Ala Glu Val Gly Gly Ser
 85 90 95
 Pro Glu Val Ser Ser
 100

<210> 397
 <211> 22
 <212> PRT
 <213> Homo sapiens

<400> 397
 Thr Thr Cys Phe Lys Lys Tyr Glu Phe Ser Phe Phe Val Leu Gly Phe
 1 5 10 15
 Leu Arg Arg Trp Gly Ala
 20

<210> 398
 <211> 26
 <212> PRT
 <213> Homo sapiens

<400> 398
 Ser Glu Lys Glu Gly Lys Asp Phe Ser Gly Phe Ala Leu Ser Ile His
 1 5 10 15
 Gly Pro Glu Arg Lys Lys Glu Glu Gly Trp
 20 25

<210> 399
 <211> 86
 <212> PRT

<213> Homo sapiens

<400> 399

Met Asn Glu Cys Ile Ala Lys Pro Cys Met Ala Ala Phe Cys Ser Cys
1 5 10 15

Pro Ser Cys Cys Leu Pro Ser Arg Pro Gly Cys Ser Arg Glu Gln Arg
20 25 30

Cys Ala Phe Ser Cys Glu Pro Cys His Thr Val Glu His Trp Val Glu
35 40 45

Pro Met Gly Gln Gly Gln Arg Gln Glu His Thr Gln Gly Ser Val Leu
50 55 60

Pro Ser Ser His Pro Ser Arg Gly Lys Ala Thr Thr Val His Ser Cys
65 70 75 80

Cys Gln Glu Pro Trp Gly
85

<210> 400

<211> 27

<212> PRT

<213> Homo sapiens

<400> 400

Phe Cys Ser Cys Pro Ser Cys Cys Leu Pro Ser Arg Pro Gly Cys Ser
1 5 10 15

Arg Glu Gln Arg Cys Ala Phe Ser Cys Glu Pro
20 25

<210> 401

<211> 23

<212> PRT

<213> Homo sapiens

<400> 401

Gly Gln Arg Gln Glu His Thr Gln Gly Ser Val Leu Pro Ser Ser His
1 5 10 15

Pro Ser Arg Gly Lys Ala Thr
20

<210> 402

<211> 139

<212> PRT

<213> Homo sapiens

<400> 402

Gly Val Val Asn Ser Cys Leu Leu Pro Leu Pro Pro Arg Leu Leu Ala
1 5 10 15

196

Thr Gly Met Asp Cys Gly Gly Phe Ala Ser Arg Arg Met Gly Gly Arg
 20 25 30
 Gln His Ala Ala Leu Ser Val Phe Leu Pro Leu Pro Leu Ala His Gly
 35 40 45
 Leu Tyr Pro Met Phe Asn Cys Val Ala Gly Leu Thr Gly Lys Gly Thr
 50 55 60
 Ser Leu Leu Ser Gly Ala Ala Arg Pro Ala Gly Glu Ala Ala Ala Arg
 65 70 75 80
 Ala Gly Thr Lys Gly Ser His Ala Arg Phe Gly Asn Ala Phe Ile His
 85 90 95
 Ser Phe Ile His Ser Phe Ile Glu Cys Leu Leu Asn Thr Tyr Cys Val
 100 105 110
 Pro Ser Ser Ala Leu Thr Ala Val Gly Ile Gly Asp Ile Leu Lys Asn
 115 120 125
 Lys Asn Asp Lys Ser Ser Cys Leu Cys Ser Cys
 130 135

<210> 403

<211> 25

<212> PRT

<213> Homo sapiens

<400> 403

Gly Met Asp Cys Gly Gly Phe Ala Ser Arg Arg Met Gly Gly Arg Gln
 1 5 10 15
 His Ala Ala Leu Ser Val Phe Leu Pro
 20 25

<210> 404

<211> 25

<212> PRT

<213> Homo sapiens

<400> 404

Leu Thr Gly Lys Gly Thr Ser Leu Leu Ser Gly Ala Ala Arg Pro Ala
 1 5 10 15
 Gly Glu Ala Ala Ala Arg Ala Gly Thr
 20 25

<210> 405

<211> 22

<212> PRT

<213> Homo sapiens

<400> 405

197

Leu Asn Thr Tyr Cys Val Pro Ser Ser Ala Leu Thr Ala Val Gly Ile
 1 5 10 15

Gly Asp Ile Leu Lys Asn
 20

<210> 406

<211> 55

<212> PRT

<213> Homo sapiens

<400> 406

Thr Ser Leu Ser Gln Leu Trp His Phe Cys His Phe Trp Pro Val Lys
 1 5 10 15

Phe Cys Cys Gly Gly Cys Pro Val His Cys Arg Met Phe Ser Ser Ile
 20 25 30

Ser Gly Leu Tyr Leu Leu Asn Ala Ser Ala Pro Ser Leu Gln Leu Asn
 35 40 45

Asp Pro Lys Cys Leu Gln Thr
 50 55

<210> 407

<211> 28

<212> PRT

<213> Homo sapiens

<400> 407

Trp Pro Val Lys Phe Cys Cys Gly Gly Cys Pro Val His Cys Arg Met
 1 5 10 15

Phe Ser Ser Ile Ser Gly Leu Tyr Leu Leu Asn Ala
 20 25

<210> 408

<211> 20

<212> PRT

<213> Homo sapiens

<400> 408

Ser Cys Arg Cys Trp Ala Leu Gly Ala Gly Gly Gly Gln Arg Gln Trp
 1 5 10 15

Val Gly Arg Ser
 20

<210> 409

<211> 80

<212> PRT

<213> Homo sapiens

198

<400> 409

Thr Gly Ala Gln Ala Pro Lys Met Gly Ala Arg Gln Arg Lys Arg Pro
 1 5 10 15

Leu Gln Thr Arg Ile Lys Asn Ser Ser Lys Ser Thr Leu Trp Pro Pro
 20 25 30

Gln Trp Val Arg Cys Gly Arg Trp Trp Thr Trp Pro Ser Arg Lys Lys
 35 40 45

Thr Ser Arg Pro Arg Arg Gln Leu Phe Thr Ser Thr Leu Ser Thr Ser
 50 55 60

Ala Ser Ala Leu Val Trp Pro Val Ser Trp Phe Ser Gln Glu Gly His
 65 70 75 80

<210> 410

<211> 25

<212> PRT

<213> Homo sapiens

<400> 410

Met Gly Ala Arg Gln Arg Lys Arg Pro Leu Gln Thr Arg Ile Lys Asn
 1 5 10 15

Ser Ser Lys Ser Thr Leu Trp Pro Pro
 20 25

<210> 411

<211> 23

<212> PRT

<213> Homo sapiens

<400> 411

Pro Arg Arg Gln Leu Phe Thr Ser Thr Leu Ser Thr Ser Ala Ser Ala
 1 5 10 15

Leu Val Trp Pro Val Ser Trp
 20

<210> 412

<211> 25

<212> PRT

<213> Homo sapiens

<400> 412

Asp Gly Gly Gly Lys Glu Glu Gly Val Ser Cys Leu Lys Ile Ser Leu
 1 5 10 15

Leu Cys Gly Pro Trp Leu Trp Leu Pro
 20 25

<210> 413
 <211> 135
 <212> PRT
 <213> Homo sapiens

<400> 413
 His Glu Met Gly Glu Leu Ala Ile Cys His Thr Arg Val Pro Phe Ser
 1 5 10 15
 Leu Pro Ser Ser Ala Gln Gly Val Pro Gln Asn Leu Gln Gly Pro Ile
 20 25 30
 Gly His Leu Ala Val Cys Thr Pro Ser Ser Leu Thr Ser Trp His Phe
 35 40 45
 Pro Gln Lys Arg Glu Lys Trp Ser Thr Val Asn Lys Arg Gln Arg Phe
 50 55 60
 Leu Gln Phe Pro Ala Pro Leu Arg Asn Trp Ile Pro Gln Thr Pro Leu
 65 70 75 80
 Ser Leu Ser Val Ser Ser Gly Pro Leu Gly Ser Phe Thr Val Phe Thr
 85 90 95
 Leu Leu Ser Leu Cys Ala Trp Pro Trp Cys Cys Arg Asp Cys Tyr Lys
 100 105 110
 Ser Cys Cys Pro Ile Pro Ile Phe Asn Leu Thr Ala Pro Leu Cys Val
 115 120 125
 His Thr Pro Glu Pro Ser Ser
 130 135

<210> 414
 <211> 23
 <212> PRT
 <213> Homo sapiens

<400> 414
 Ser Ser Ala Gln Gly Val Pro Gln Asn Leu Gln Gly Pro Ile Gly His
 1 5 10 15
 Leu Ala Val Cys Thr Pro Ser
 20

<210> 415
 <211> 28
 <212> PRT
 <213> Homo sapiens

<400> 415
 Val Asn Lys Arg Gln Arg Phe Leu Gln Phe Pro Ala Pro Leu Arg Asn
 1 5 10 15

Trp Ile Pro Gln Thr Pro Leu Ser Leu Ser Val Ser
20 25

<210> 416
<211> 23
<212> PRT
<213> Homo sapiens

<400> 416
Cys Cys Arg Asp Cys Tyr Lys Ser Cys Cys Pro Ile Pro Ile Phe Asn
1 5 10 15

Leu Thr Ala Pro Leu Cys Val
20

<210> 417
<211> 150
<212> PRT
<213> Homo sapiens

<400> 417
Asp Leu Asn Val Thr Asn Glu Gly Glu Gly Lys Glu Val Leu Gly Gln
1 5 10 15

Gly Ser Thr Asn Asn Glu Lys Lys Cys Gln Lys Ala Thr Ser Asn Thr
20 25 30

Glu Pro Arg Ala Arg Glu Ala Lys Ala Arg His Ala Asn Met Gly Thr
35 40 45

Ser Asp Arg Glu Ser Pro Thr Trp Ser Leu Thr Ala Glu Gly Leu Lys
50 55 60

Ala Lys Ser Lys Met Gln Gly Lys Ala Thr Lys Gly Ala Ala Ser Thr
65 70 75 80

Met Gly Ser His Asn Gln Gly Pro His Lys Arg Glu Ile Phe Lys His
85 90 95

Glu Thr Pro Ser Ser Phe Pro Pro Pro Ser Gln Cys Gln Pro Glu Leu
100 105 110

Leu Pro Tyr Lys Tyr Trp Ala Thr Leu Ala Ser Gly Tyr Val Pro Ser
115 120 125

Trp Leu Pro Ser Val Asp Ser Tyr Arg Ile Asn Thr Ala Ile Lys Asp
130 135 140

Lys Asn Gly Gln Asp Thr
145 150

<210> 418
<211> 24

<212> PRT

<213> Homo sapiens

<400> 418

Val Leu Gly Gln Gly Ser Thr Asn Asn Glu Lys Lys Cys Gln Lys Ala
1 5 10 15

Thr Ser Asn Thr Glu Pro Arg Ala
20

<210> 419

<211> 29

<212> PRT

<213> Homo sapiens

<400> 419

Arg Glu Ser Pro Thr Trp Ser Leu Thr Ala Glu Gly Leu Lys Ala Lys
1 5 10 15

Ser Lys Met Gln Gly Lys Ala Thr Lys Gly Ala Ala Ser
20 25

<210> 420

<211> 22

<212> PRT

<213> Homo sapiens

<400> 420

Gly Tyr Val Pro Ser Trp Leu Pro Ser Val Asp Ser Tyr Arg Ile Asn
1 5 10 15

Thr Ala Ile Lys Asp Lys
20

<210> 421

<211> 12

<212> PRT

<213> Homo sapiens

<400> 421

Asn Ser Ala Glu Gln Ser Met Leu Ile Leu Val Thr
1 5 10

<210> 422

<211> 122

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (2)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (5)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 422

Arg	Xaa	Asp	Arg	Xaa	Pro	Val	Pro	Glu	Leu	Pro	Gly	Tyr	Glu	Pro	Thr
1				5					10					15	

Arg	Thr	Asp	Ile	Ser	Ser	Phe	Lys	Asn	Ile	Tyr	Arg	Tyr	Ala	Phe	Asp
		20						25					30		

Phe	Ala	Arg	Asp	Lys	Asp	Gln	Arg	Ser	Leu	Asp	Ile	Asp	Thr	Ala	Lys
		35					40					45			

Ser	Met	Leu	Ala	Leu	Leu	Leu	Gly	Arg	Thr	Trp	Pro	Leu	Phe	Ser	Val
	50					55					60				

Phe	Tyr	Gln	Tyr	Leu	Glu	Gln	Ser	Lys	Tyr	Arg	Val	Met	Asn	Lys	Asp
65					70					75					80

Gln	Trp	Tyr	Asn	Val	Leu	Glu	Phe	Ser	Arg	Thr	Val	His	Ala	Asp	Leu
			85						90					95	

Ser	Asn	Tyr	Asp	Glu	Asp	Gly	Ala	Trp	Pro	Val	Leu	Leu	Asp	Glu	Phe
			100					105					110		

Val	Glu	Trp	Gln	Lys	Val	Arg	Gln	Thr	Ser
	115						120		

<210> 423

<211> 28

<212> PRT

<213> Homo sapiens

<400> 423

Pro	Thr	Arg	Thr	Asp	Ile	Ser	Ser	Phe	Lys	Asn	Ile	Tyr	Arg	Tyr	Ala
1				5					10					15	

Phe	Asp	Phe	Ala	Arg	Asp	Lys	Asp	Gln	Arg	Ser	Leu
		20						25			

<210> 424

<211> 29

<212> PRT

<213> Homo sapiens

<400> 424

Ser	Met	Leu	Ala	Leu	Leu	Leu	Gly	Arg	Thr	Trp	Pro	Leu	Phe	Ser	Val
1				5					10					15	

Phe	Tyr	Gln	Tyr	Leu	Glu	Gln	Ser	Lys	Tyr	Arg	Val	Met
		20						25				

203

<210> 425
<211> 27
<212> PRT
<213> Homo sapiens

<400> 425
Phe Ser Arg Thr Val His Ala Asp Leu Ser Asn Tyr Asp Glu Asp Gly
1 5 10 15
Ala Trp Pro Val Leu Leu Asp Glu Phe Val Glu
20 25

<210> 426
<211> 10
<212> PRT
<213> Homo sapiens

<400> 426
Ile Tyr Arg Tyr Ala Phe Asp Phe Ala Arg
1 5 10

<210> 427
<211> 8
<212> PRT
<213> Homo sapiens

<400> 427
Lys Asp Gln Arg Ser Leu Asp Ile
1 5

<210> 428
<211> 8
<212> PRT
<213> Homo sapiens

<400> 428
Asn Val Leu Glu Phe Ser Arg Thr
1 5

<210> 429
<211> 21
<212> PRT
<213> Homo sapiens

<400> 429
Asp Leu Ser Asn Tyr Asp Glu Asp Gly Ala Trp Pro Val Leu Leu Asp
1 5 10 15
Glu Phe Val Glu Trp
20

<210> 430

<211> 37
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (15)
 <223> Xaa equals any of the naturally occurring L-amino acids

<400> 430
 Leu Phe Arg Cys Pro Ile Gly Lys Ala Gly Thr Pro Ala Gly Xaa Gly
 1 5 10 15
 Pro Glu Phe Pro Gly Arg Pro Thr Arg Pro Val Arg Glu Lys Glu Leu
 20 25 30
 Thr Glu Thr Phe Glu
 35

<210> 431
 <211> 21
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (9)
 <223> Xaa equals any of the naturally occurring L-amino acids

<400> 431
 Gly Lys Ala Gly Thr Pro Ala Gly Xaa Gly Pro Glu Phe Pro Gly Arg
 1 5 10 15
 Pro Thr Arg Pro Val
 20

<210> 432
 <211> 45
 <212> PRT
 <213> Homo sapiens

<400> 432
 Phe Phe Val Phe Pro Tyr Pro Tyr Pro Phe Arg Pro Leu Pro Pro Ile
 1 5 10 15
 Pro Phe Pro Arg Phe Pro Trp Phe Arg Arg Asn Phe Pro Ile Pro Ile
 20 25 30
 Pro Glu Ser Ala Pro Thr Thr Pro Leu Pro Ser Glu Lys
 35 40 45

<210> 433
 <211> 21
 <212> PRT

<213> Homo sapiens

<400> 433

Pro Trp Phe Arg Arg Asn Phe Pro Ile Pro Ile Pro Glu Ser Ala Pro
 1 5 10 15

Thr Thr Pro Leu Pro
 20

<210> 434

<211> 61

<212> PRT

<213> Homo sapiens

<400> 434

Phe Tyr Pro Pro Met Thr Gln Gly Lys Glu Ser Leu Pro Leu Leu Ala
 1 5 10 15

Leu Gln Ile Phe Asn Thr Thr Phe Arg Pro Ser Phe Ala Phe Phe Ser
 20 25 30

Gly His Arg Thr Leu Phe Phe Gly Val Arg Ser Pro Asn Pro Pro Lys
 35 40 45

Pro Arg Ile Phe Leu Ile Trp Leu Ile Ala Val Ala Leu
 50 55 60

<210> 435

<211> 31

<212> PRT

<213> Homo sapiens

<400> 435

Leu Leu Ala Leu Gln Ile Phe Asn Thr Thr Phe Arg Pro Ser Phe Ala
 1 5 10 15

Phe Phe Ser Gly His Arg Thr Leu Phe Phe Gly Val Arg Ser Pro
 20 25 30

<210> 436

<211> 52

<212> PRT

<213> Homo sapiens

<400> 436

His Leu Ala Gln Thr Val Met Met His Pro Gln Lys Ser Phe Tyr Gln
 1 5 10 15

Val Lys Asn Thr Asn His Ser Asp Arg Gly Ala Ile Glu Glu Thr Gln
 20 25 30

Ile Leu Glu Asp Arg Leu Gly Gln Ile Pro Leu Cys Leu Glu Ser Gln
 35 40 45

Ile Trp Glu Ala
50

<210> 437
<211> 28
<212> PRT
<213> Homo sapiens

<400> 437
Lys Asn Thr Asn His Ser Asp Arg Gly Ala Ile Glu Glu Thr Gln Ile
1 5 10 15
Leu Glu Asp Arg Leu Gly Gln Ile Pro Leu Cys Leu
20 25

<210> 438
<211> 73
<212> PRT
<213> Homo sapiens

<400> 438
Gln Gly Cys Tyr Arg Arg Asp Ser Asn Ile Gly Arg Gln Val Arg Pro
1 5 10 15
Asp Ser Ile Met Leu Arg Lys Pro Asp Leu Gly Ser Ile Thr His Tyr
20 25 30
Gly Ser Val Leu Gly Asn Leu Asn Tyr Cys Asp Leu Pro Gln Leu Tyr
35 40 45
Arg Asn Pro Ser Leu Gly Asn Ser Gly Met Arg Glu Met Phe Ser Pro
50 55 60
Phe Tyr Asn Pro Val Glu Cys His Pro
65 70

<210> 439
<211> 23
<212> PRT
<213> Homo sapiens

<400> 439
Pro Asp Ser Ile Met Leu Arg Lys Pro Asp Leu Gly Ser Ile Thr His
1 5 10 15
Tyr Gly Ser Val Leu Gly Asn
20

<210> 440
<211> 22
<212> PRT
<213> Homo sapiens

207

<400> 440

Tyr Arg Asn Pro Ser Leu Gly Asn Ser Gly Met Arg Glu Met Phe Ser
 1 5 10 15

Pro Phe Tyr Asn Pro Val
 20

<210> 441

<211> 21

<212> PRT

<213> Homo sapiens

<400> 441

Asn Ser Ala Arg Gly Leu Ser Gly Gly His Pro Phe Pro Trp Leu Ser
 1 5 10 15

Glu Gly His Pro Phe
 20

<210> 442

<211> 107

<212> PRT

<213> Homo sapiens

<400> 442

Thr Asp Ser Asp Leu Thr Leu Gly Ile Leu Leu Leu Gly Ile Tyr Thr
 1 5 10 15

Asn His Ile Trp Glu Met Phe Leu Ala Ala Ser Arg Ile Asn Ser Pro
 20 25 30

Lys Leu Glu Pro Glu Lys Ser Val Lys Arg Gln Ile Asn Phe Pro Ser
 35 40 45

Ser Lys Asp Val Gly Cys Ser Leu Glu Val Pro Lys Asp Gly Pro Pro
 50 55 60

Leu Ser His Gly Lys Glu Trp Ile Pro Leu Ser His Arg Lys Gly Trp
 65 70 75 80

Ile Pro Leu Ser His Met Lys Gly Trp Pro Ser Leu Ser His Gly Lys
 85 90 95

Gly Trp Pro Pro Leu Ser Pro Arg Ala Glu Phe
 100 105

<210> 443

<211> 20

<212> PRT

<213> Homo sapiens

<400> 443

Leu Gly Ile Leu Leu Leu Gly Ile Tyr Thr Asn His Ile Trp Glu Met
 1 5 10 15

Phe Leu Ala Ala
20

<210> 444
<211> 27
<212> PRT
<213> Homo sapiens

<400> 444
Lys Ser Val Lys Arg Gln Ile Asn Phe Pro Ser Ser Lys Asp Val Gly
1 5 10 15

Cys Ser Leu Glu Val Pro Lys Asp Gly Pro Pro
20 25

<210> 445
<211> 27
<212> PRT
<213> Homo sapiens

<400> 445
Gly Lys Glu Trp Ile Pro Leu Ser His Arg Lys Gly Trp Ile Pro Leu
1 5 10 15

Ser His Met Lys Gly Trp Pro Ser Leu Ser His
20 25

<210> 446
<211> 47
<212> PRT
<213> Homo sapiens

<400> 446
Gly Trp Ala Ser Thr Gln Pro Arg Glu Arg Met Asp Pro Ala Gln Pro
1 5 10 15

Gln Glu Arg Met Asp Pro Ser Gln Pro His Glu Arg Met Ala Leu Thr
20 25 30

Gln Pro Trp Lys Arg Met Ala Pro Thr Gln Pro Ser Cys Arg Ile
35 40 45

<210> 447
<211> 24
<212> PRT
<213> Homo sapiens

<400> 447
Pro Ala Gln Pro Gln Glu Arg Met Asp Pro Ser Gln Pro His Glu Arg
1 5 10 15

Met Ala Leu Thr Gln Pro Trp Lys

20

<210> 448
<211> 30
<212> PRT
<213> Homo sapiens

<400> 448
Ile Ala Asn Gly Gly Gly Arg Pro Ile Lys Leu Asn Ala Leu Tyr Lys
1 5 10 15
Ile Gln Asn Glu Cys Lys Ile Val Phe Thr Cys Ile Asp Phe
20 25 30

<210> 449
<211> 33
<212> PRT
<213> Homo sapiens

<400> 449
Met Pro Cys Ile Lys Ser Lys Met Asn Ala Lys Leu Phe Ser Leu Val
1 5 10 15
Leu Thr Leu Cys Cys Met Ile Pro Ile Ser Val Leu Phe Gly Thr Cys
20 25 30

Ile

<210> 450
<211> 101
<212> PRT
<213> Homo sapiens

<400> 450
Gln Val Ala Met Gly Ser Leu Ser Gly Leu Arg Leu Ala Ala Gly Ser
1 5 10 15
Cys Phe Arg Leu Cys Glu Arg Asp Val Ser Ser Ser Leu Arg Leu Thr
20 25 30
Arg Ser Ser Asp Leu Lys Arg Ile Asn Gly Phe Cys Thr Lys Pro Gln
35 40 45
Glu Ser Pro Gly Ala Pro Ser Arg Thr Tyr Asn Arg Val Pro Leu His
50 55 60
Lys Pro Thr Asp Trp Gln Lys Lys Ile Leu Ile Trp Ser Gly Arg Phe
65 70 75 80
Lys Lys Glu Asp Glu Ile Pro Glu Thr Val Ser Leu Glu Met Leu Asp
85 90 95
Ala Ala Lys Asn Lys

210

100

<210> 451
<211> 25
<212> PRT
<213> Homo sapiens

<400> 451
Gly Leu Arg Leu Ala Ala Gly Ser Cys Phe Arg Leu Cys Glu Arg Asp
1 5 10 15

Val Ser Ser Ser Leu Arg Leu Thr Arg
20 25

<210> 452
<211> 20
<212> PRT
<213> Homo sapiens

<400> 452
Ala Pro Ser Arg Thr Tyr Asn Arg Val Pro Leu His Lys Pro Thr Asp
1 5 10 15

Trp Gln Lys Lys
20

<210> 453
<211> 23
<212> PRT
<213> Homo sapiens

<400> 453
Ile Trp Ser Gly Arg Phe Lys Lys Glu Asp Glu Ile Pro Glu Thr Val
1 5 10 15

Ser Leu Glu Met Leu Asp Ala
20

<210> 454
<211> 63
<212> PRT
<213> Homo sapiens

<400> 454
Met Asp Phe Ala Gln Asn His Arg Lys Val Pro Glu Leu His Pro Ala
1 5 10 15

Leu Thr Thr Glu Cys Leu Tyr Thr Asn Leu Arg Ile Gly Arg Lys Arg
20 25 30

Ser Ser Tyr Gly Gln Val Ala Ser Lys Arg Lys Met Lys Ser Gln Arg
35 40 45

211

Leu Ser Arg Trp Arg Cys Leu Met Leu Gln Arg Thr Arg Cys Glu
 50 55 60

<210> 455
 <211> 19
 <212> PRT
 <213> Homo sapiens

<400> 455
 Lys Val Pro Glu Leu His Pro Ala Leu Thr Thr Glu Cys Leu Tyr Thr
 1 5 10 15

Asn Leu Arg

<210> 456
 <211> 26
 <212> PRT
 <213> Homo sapiens

<400> 456
 Lys Arg Ser Ser Tyr Gly Gln Val Ala Ser Lys Arg Lys Met Lys Ser
 1 5 10 15

Gln Arg Leu Ser Arg Trp Arg Cys Leu Met
 20 25

<210> 457
 <211> 12
 <212> PRT
 <213> Homo sapiens

<400> 457
 Ile Asn Gly Phe Cys Thr Lys Pro Gln Glu Ser Pro
 1 5 10

<210> 458
 <211> 9
 <212> PRT
 <213> Homo sapiens

<400> 458
 Arg Val Pro Leu His Lys Pro Thr Asp
 1 5

<210> 459
 <211> 8
 <212> PRT
 <213> Homo sapiens

<400> 459
 Trp Ser Gly Arg Phe Lys Lys Glu

1 5

<210> 460
<211> 9
<212> PRT
<213> Homo sapiens

<400> 460
Glu Met Leu Asp Ala Ala Lys Asn Lys
1 5

<210> 461
<211> 9
<212> PRT
<213> Homo sapiens

<400> 461
Ser Tyr Leu Met Ile Ala Leu Thr Val
1 5

<210> 462
<211> 9
<212> PRT
<213> Homo sapiens

<400> 462
Met Val Ile Glu Gly Lys Lys Ala Ala
1 5

<210> 463
<211> 68
<212> PRT
<213> Homo sapiens

<400> 463
Arg Pro Gly Met Arg Ala Leu Gly Ser Cys Leu Ser Leu Leu Ala Leu
1 5 10 15

Cys Ser Pro Gln Ala Arg Pro Gly Pro Arg Thr Leu Asp Ala Ser Thr
20 25 30

Ala Thr Leu Thr Pro His Phe Ser Pro Cys Ala Arg Phe Ser Pro Val
35 40 45

Gly Pro Ser Ala Val Pro Phe Ala Ala Thr Pro Leu Pro Leu Ala Gly
50 55 60

Pro His Gln Pro
65

<210> 464
<211> 20

213

<212> PRT

<213> Homo sapiens

<400> 464

Gly Ser Cys Leu Ser Leu Leu Ala Leu Cys Ser Pro Gln Ala Arg Pro
 1 5 10 15

Gly Pro Arg Thr
 20

<210> 465

<211> 23

<212> PRT

<213> Homo sapiens

<400> 465

His Phe Ser Pro Cys Ala Arg Phe Ser Pro Val Gly Pro Ser Ala Val
 1 5 10 15

Pro Phe Ala Ala Thr Pro Leu
 20

<210> 466

<211> 92

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (43)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (80)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 466

Ala Ile Glu Glu Arg Asn Lys Ser Arg Leu Thr Gln Gln Ala Ser Glu
 1 5 10 15

Pro Thr Gly Ser Pro Arg Tyr Leu His Glu Gln His Pro Gly Ser Arg
 20 25 30

Ser Gln Met Asp Cys Gly Ser Leu Thr Met Xaa Cys Pro Pro Pro Arg
 35 40 45

Val Arg Asp Asp Arg Thr Ser Ala Arg Gly Val Pro Arg Gln Ala Ala
 50 55 60

Pro Asp Ile Val Gly Gly Arg Pro Ser Ser Arg Ala Cys Val Ser Xaa
 65 70 75 80

Pro Ala Cys Ala Pro Ser Ala Ala Val Phe Pro Tyr
 85 90

SUBSTITUTE SHEET (RULE 26)

<210> 467
 <211> 24
 <212> PRT
 <213> Homo sapiens

<400> 467
 Leu Thr Gln Gln Ala Ser Glu Pro Thr Gly Ser Pro Arg Tyr Leu His
 1 5 10 15

Glu Gln His Pro Gly Ser Arg Ser
 20

<210> 468
 <211> 25
 <212> PRT
 <213> Homo sapiens

<400> 468
 Ser Ala Arg Gly Val Pro Arg Gln Ala Ala Pro Asp Ile Val Gly Gly
 1 5 10 15

Arg Pro Ser Ser Arg Ala Cys Val Ser
 20 25

<210> 469
 <211> 14
 <212> PRT
 <213> Homo sapiens

<400> 469
 Pro Arg Val Arg Lys Thr Pro His Leu Ser Ala Ser Gly Lys
 1 5 10

<210> 470
 <211> 59
 <212> PRT
 <213> Homo sapiens

<400> 470
 Tyr Tyr Tyr Ser Met Leu Lys Ile Cys His Ile Thr Ile Leu Glu Thr
 1 5 10 15

Leu Ser Asp Arg Thr Pro Arg Lys Phe Ala Lys Lys Cys Tyr Ile Leu
 20 25 30

Tyr Ile Lys Leu Ser Asp Ser Ser Val Glu Lys Val Ala Tyr Thr Leu
 35 40 45

Leu Leu Leu Ile Pro Ala Ala Ile Glu Lys Lys
 50 55

215

<210> 471
 <211> 32
 <212> PRT
 <213> Homo sapiens

<400> 471
 Thr Ile Leu Glu Thr Leu Ser Asp Arg Thr Pro Arg Lys Phe Ala Lys
 1 5 10 15
 Lys Cys Tyr Ile Leu Tyr Ile Lys Leu Ser Asp Ser Ser Val Glu Lys
 20 25 30

<210> 472
 <211> 17
 <212> PRT
 <213> Homo sapiens

<400> 472
 Val His Thr Lys Glu Ile Phe Arg Glu Arg Ser Ala Gly Phe Pro Val
 1 5 10 15
 Lys

<210> 473
 <211> 97
 <212> PRT
 <213> Homo sapiens

<400> 473
 Leu Glu Met Gly Phe Gln Pro Thr Lys Glu Ile Asn Ala Arg Gly Ser
 1 5 10 15
 Glu Pro Cys Gln Ala Gln Ser Thr Ser Leu Pro Lys Leu Pro Arg Trp
 20 25 30
 Gly Ser Arg Pro Glu Ala Pro Gln Thr Pro Gln Gly Gly Leu Glu Ser
 35 40 45
 Arg Cys Cys Thr Pro Val Ser Lys Gln Ser Leu Asn Leu Lys Ala Asp
 50 55 60
 Arg Phe Lys Ala Leu Thr Leu Gly Arg Ala Gln Trp Leu Thr Pro Val
 65 70 75 80
 Ile Gln Ala Leu Ser Glu Leu Arg Trp Val Asp His Leu Arg Ser Gly
 85 90 95
 Val

216

<210> 474
 <211> 24
 <212> PRT
 <213> Homo sapiens

<400> 474
 Phe Gln Pro Thr Lys Glu Ile Asn Ala Arg Gly Ser Glu Pro Cys Gln
 1 5 10 15

Ala Gln Ser Thr Ser Leu Pro Lys
 20

<210> 475
 <211> 27
 <212> PRT
 <213> Homo sapiens

<400> 475
 Pro Lys Leu Pro Arg Trp Gly Ser Arg Pro Glu Ala Pro Gln Thr Pro
 1 5 10 15

Gln Gly Gly Leu Glu Ser Arg Cys Cys Thr Pro
 20 25

<210> 476
 <211> 27
 <212> PRT
 <213> Homo sapiens

<400> 476
 Arg Phe Lys Ala Leu Thr Leu Gly Arg Ala Gln Trp Leu Thr Pro Val
 1 5 10 15

Ile Gln Ala Leu Ser Glu Leu Arg Trp Val Asp
 20 25

<210> 477
 <211> 176
 <212> PRT
 <213> Homo sapiens

<400> 477
 Arg Ile Pro Leu Gln Ser Asp Gly Ser Phe Leu His Glu Lys Ser Ser
 1 5 10 15

Gln Gln Arg Ser Asn Arg Asn Phe Pro Cys Pro Thr Leu Gln Cys Asn
 20 25 30

Pro Glu Val Ser Phe Trp Phe Val Val Thr Asp Pro Ser Lys Asn His
 35 40 45

Thr Leu Pro Ala Val Glu Val Gln Ser Ala Ile Arg Met Asn Lys Asn
 50 55 60

217

Arg Ile Asn Asn Ala Phe Phe Leu Asn Asp Gln Thr Leu Glu Phe Leu
 65 70 75 80
 Lys Ile Pro Ser Thr Leu Ala Pro Pro Met Asp Pro Ser Val Pro Ile
 85 90 95
 Trp Ile Ile Ile Phe Gly Val Ile Phe Cys Ile Ile Ile Val Ala Ile
 100 105 110
 Ala Leu Leu Ile Leu Ser Gly Ile Trp Gln Arg Arg Arg Lys Asn Lys
 115 120 125
 Glu Pro Ser Glu Val Asp Asp Ala Glu Asp Lys Cys Glu Asn Met Ile
 130 135 140
 Thr Ile Glu Asn Gly Ile Pro Ser Asp Pro Leu Asp Met Lys Gly Gly
 145 150 155 160
 His Ile Asn Asp Ala Phe Met Thr Glu Asp Glu Arg Leu Thr Pro Leu
 165 170 175

<210> 478
 <211> 25
 <212> PRT
 <213> Homo sapiens

<400> 478
 Pro Cys Pro Thr Leu Gln Cys Asn Pro Glu Val Ser Phe Trp Phe Val
 1 5 10 15
 Val Thr Asp Pro Ser Lys Asn His Thr
 20 25

<210> 479
 <211> 23
 <212> PRT
 <213> Homo sapiens

<400> 479
 Ala Ile Arg Met Asn Lys Asn Arg Ile Asn Asn Ala Phe Phe Leu Asn
 1 5 10 15
 Asp Gln Thr Leu Glu Phe Leu
 20

<210> 480
 <211> 24
 <212> PRT
 <213> Homo sapiens
 <400> 480

218

Ile Trp Gln Arg Arg Arg Lys Asn Lys Glu Pro Ser Glu Val Asp Asp
 1 5 10 15

Ala Glu Asp Lys Cys Glu Asn Met
 20

<210> 481

<211> 19

<212> PRT

<213> Homo sapiens

<400> 481

Pro Leu Asp Met Lys Gly Gly His Ile Asn Asp Ala Phe Met Thr Glu
 1 5 10 15

Asp Glu Arg

<210> 482

<211> 136

<212> PRT

<213> Homo sapiens

<400> 482

Gly Ser Arg Thr Thr Ala Leu Gln Arg Gly Val Ser Leu Ser Ser Ser
 1 5 10 15

Val Met Lys Ala Ser Leu Ile Cys Pro Pro Phe Met Ser Arg Gly Ser
 20 25 30

Glu Gly Met Pro Phe Ser Ile Val Ile Met Phe Ser His Leu Ser Ser
 35 40 45

Ala Ser Ser Thr Ser Asp Gly Ser Leu Phe Phe Leu Leu Arg Cys Gln
 50 55 60

Ile Pro Asp Lys Ile Ser Ser Ala Ile Ala Thr Met Met Met Gln Asn
 65 70 75 80

Ile Thr Pro Asn Ile Ile Ile Gln Met Gly Thr Asp Gly Ser Met Gly
 85 90 95

Gly Ala Ser Val Glu Gly Ile Phe Lys Asn Ser Arg Val Trp Ser Phe
 100 105 110

Arg Lys Lys Ala Leu Leu Ile Arg Phe Leu Phe Ile Leu Met Ala Asp
 115 120 125

Cys Thr Ser Thr Ala Gly Arg Val
 130 135

<210> 483

<211> 28

<212> PRT

219

<213> Homo sapiens

<400> 483

Val Ser Leu Ser Ser Val Met Lys Ala Ser Leu Ile Cys Pro Pro
 1 5 10 15

Phe Met Ser Arg Gly Ser Glu Gly Met Pro Phe Ser
 20 25

<210> 484

<211> 24

<212> PRT

<213> Homo sapiens

<400> 484

Ser Met Gly Gly Ala Ser Val Glu Gly Ile Phe Lys Asn Ser Arg Val
 1 5 10 15

Trp Ser Phe Arg Lys Lys Ala Leu
 20

<210> 485

<211> 29

<212> PRT

<213> Homo sapiens

<400> 485

Gly Ala Arg Gly Ser Gln Gln Asp Ala Pro Ala Leu Gln Glu Ala Glu
 1 5 10 15

Val Arg Gly Pro Glu Arg Ala Gln Pro Ala Arg Gly Arg
 20 25

<210> 486

<211> 439

<212> PRT

<213> Homo sapiens

<400> 486

Ser Glu Arg Pro Gly Glu Gly Pro Ala Arg Pro Gly Gln Asp Asp Gln
 1 5 10 15

Gly Pro Ala Val Pro Ala Val Ala Gly Ala Gly Val Gly Val His Asp
 20 25 30

Pro Ala Asp His Arg Val Leu Gly Gln Arg Ser Ala Ala His Phe Tyr
 35 40 45

Leu His Thr Ser Phe Ser Arg Pro His Thr Gly Pro Pro Leu Pro Thr
 50 55 60

Pro Gly Pro Asp Arg Thr Gly Ser Ser Arg Pro Thr Pro Met Ser Thr
 65 70 75 80

SUBSTITUTE SHEET (RULE 26)

220

Ser Phe Trp Thr Ile Ser His Ala Gly Val Lys Gln Ser Asp Leu Pro
85 90 95

Arg Lys Glu Thr Glu Gln Pro Pro Ala Pro Gly Glu His Gly Gly Glu
100 105 110

Arg Glu Arg Leu Arg Leu Val Pro Ala Arg Arg Pro Ala Gln Pro Arg
115 120 125

Pro Gly Pro Ala Ala Gly Gly Ala Glu Glu Arg Ala Ala Gly Leu Leu
130 135 140

Arg Gln Leu Gln Pro Gly Leu Pro His Gln Gly Ala Arg Ile Arg Arg
145 150 155 160

His Pro Gln Leu Gly Ala Glu Pro Pro Asp Arg Gly Arg Pro Ala Arg
165 170 175

Gly His Leu Leu Leu Arg Ala Gln Gly Gly Leu His Gln Leu Glu Ala
180 185 190

Arg Asp Asp Arg Ala Glu Arg Lys Pro Ala Ala Pro Arg Cys Ala Leu
195 200 205

Pro Arg Pro Ala Ala His Pro Ala Arg Ala Arg Ala Gln Arg Gln Arg
210 215 220

Ala Pro Asp Leu Gln Gln Val Leu Ala Pro Leu Arg Glu Ala Leu Pro
225 230 235 240

Pro Pro His Glu Gly Gln Ala Gln Glu Val His Gln Val Pro Leu Arg
245 250 255

Ala Arg Pro Leu Arg Ala Pro Asp Leu Arg Leu Pro Gln Gln Val Arg
260 265 270

Ala Gly Glu Arg Gly Val Leu Pro Gln Val Arg Arg Ala His Ala Ala
275 280 285

Gly Val Arg Gln Pro His Gln Pro Ala Arg Leu Gly Ala Arg Gly Leu
290 295 300

Pro Arg Trp Pro Gln Gly Val Leu Arg Gln Leu His Pro Val Pro Ala
305 310 315 320

Gly Pro Ala His Gly Glu Ala Gly Ala Leu Gln Arg Ala Leu Ala Ala
325 330 335

Gly Val Pro Pro Leu Pro Pro Val Pro Asp Arg Leu Arg Phe Leu Gly
340 345 350

Lys Leu Glu Thr Leu Asp Glu Asp Ala Ala Gln Leu Leu Gln Leu Leu
355 360 365

Gln Val Asp Arg Gln Ser Ala Ser Pro Arg Ala Thr Gly Thr Gly Pro
370 375 380

221

Pro Ala Ala Gly Arg Arg Thr Gly Ser Pro Arg Ser Pro Trp Pro Gly
 385 390 395 400

Gly Ser Ser Cys Ile Asn Ser Thr Arg Pro Thr Leu Phe Ser Ser Ala
 405 410 415

Thr Pro Ser Pro Lys Thr Ser Ser Glu Thr Glu Ser Phe Arg Val Ala
 420 425 430

Phe Ser Arg Val Pro Gly Thr
 435

<210> 487

<211> 25

<212> PRT

<213> Homo sapiens

<400> 487

Arg Pro Gly Gln Asp Asp Gln Gly Pro Ala Val Pro Ala Val Ala Gly
 1 5 10 15

Ala Gly Val Gly Val His Asp Pro Ala
 20 25

<210> 488

<211> 21

<212> PRT

<213> Homo sapiens

<400> 488

Ser Arg Pro His Thr Gly Pro Pro Leu Pro Thr Pro Gly Pro Asp Arg
 1 5 10 15

Thr Gly Ser Ser Arg
 20

<210> 489

<211> 23

<212> PRT

<213> Homo sapiens

<400> 489

Ser His Ala Gly Val Lys Gln Ser Asp Leu Pro Arg Lys Glu Thr Glu
 1 5 10 15

Gln Pro Pro Ala Pro Gly Glu
 20

<210> 490

<211> 23

<212> PRT

<213> Homo sapiens

222

<400> 490

Arg Arg Pro Ala Gln Pro Arg Pro Gly Pro Ala Ala Gly Gly Ala Glu
 1 5 10 15

Glu Arg Ala Ala Gly Leu Leu
 20

<210> 491

<211> 23

<212> PRT

<213> Homo sapiens

<400> 491

Arg Arg His Pro Gln Leu Gly Ala Glu Pro Pro Asp Arg Gly Arg Pro
 1 5 10 15

Ala Arg Gly His Leu Leu Leu
 20

<210> 492

<211> 25

<212> PRT

<213> Homo sapiens

<400> 492

Arg Asp Asp Arg Ala Glu Arg Lys Pro Ala Ala Pro Arg Cys Ala Leu
 1 5 10 15

Pro Arg Pro Ala Ala His Pro Ala Arg
 20 25

<210> 493

<211> 27

<212> PRT

<213> Homo sapiens

<400> 493

Arg Ala Pro Asp Leu Gln Gln Val Leu Ala Pro Leu Arg Glu Ala Leu
 1 5 10 15

Pro Pro Pro His Glu Gly Gln Ala Gln Glu Val
 20 25

<210> 494

<211> 26

<212> PRT

<213> Homo sapiens

<400> 494

Asp Leu Arg Leu Pro Gln Gln Val Arg Ala Gly Glu Arg Gly Val Leu
 1 5 10 15

Pro Gln Val Arg Arg Ala His Ala Ala Gly

223

20

25

<210> 495
<211> 27
<212> PRT
<213> Homo sapiens

<400> 495
Gln Pro Ala Arg Leu Gly Ala Arg Gly Leu Pro Arg Trp Pro Gln Gly
1 5 10 15

Val Leu Arg Gln Leu His Pro Val Pro Ala Gly
20 25

<210> 496
<211> 24
<212> PRT
<213> Homo sapiens

<400> 496
Ala Gly Val Pro Pro Leu Pro Pro Val Pro Asp Arg Leu Arg Phe Leu
1 5 10 15

Gly Lys Leu Glu Thr Leu Asp Glu
20

<210> 497
<211> 25
<212> PRT
<213> Homo sapiens

<400> 497
Gln Leu Leu Gln Leu Leu Gln Val Asp Arg Gln Ser Ala Ser Pro Arg
1 5 10 15

Ala Thr Gly Thr Gly Pro Pro Ala Ala
20 25

<210> 498
<211> 25
<212> PRT
<213> Homo sapiens

<400> 498
Asn Ser Thr Arg Pro Thr Leu Phe Ser Ser Ala Thr Pro Ser Pro Lys
1 5 10 15

Thr Ser Ser Glu Thr Glu Ser Phe Arg
20 25

<210> 499
<211> 324

<212> PRT

<213> Homo sapiens

<400> 499

Leu Gly Gly Lys Arg Thr Ala Gly Pro Pro Gly Val Ala Ala Ala Ala
 1 5 10 15
 Ala Arg Arg Pro Arg Pro Glu Ser Pro Ala Ser Pro Gly Ile Val Val
 20 25 30
 Asp Leu Ala Arg Val Ala Glu Ala Val His Leu Pro Pro Val Leu Val
 35 40 45
 Glu Gly Arg Gln Leu Leu Arg Val Arg Val Gln Gln Val Leu Asp Glu
 50 55 60
 Val Gly Glu Gly His Leu Glu Ala Ser Ala Glu Gly Leu Ala Arg Arg
 65 70 75 80
 Gly Gly Gln Ala Gly Val Val Gly Val His Pro Gln His Gly His Gly
 85 90 95
 Glu Leu Ala Val Glu Leu Leu Val Leu Gln Leu Glu Leu Ala Ala Glu
 100 105 110
 Gly Gly Asp Gln Ala His Glu Gly Val Ala His Glu Glu Glu Leu Gly
 115 120 125
 Val Leu Leu Glu Leu Asp Leu His Glu Val Ala Gly Glu Leu Pro Val
 130 135 140
 Ala Ala Pro Glu Leu Val Glu Gly Gln Val Arg Ala Gly Val Val His
 145 150 155 160
 Val Leu Ala Arg Asp Ala Gln Arg Val Ala Val Gly Arg Thr Ala Val
 165 170 175
 Gln Gln Ala Ser Ala Gln His Asp His His Ala Leu Pro Val Gly Ala
 180 185 190
 Gly His Leu Gly His Val Ala Val Asp Gly Pro Val Pro Val Val His
 195 200 205
 Asp Gln Val Ala Gln Leu Arg Val Gly Asp Val Val Glu Cys Ala Leu
 210 215 220
 Leu Gly Gly Glu Gly Gln Ala Gly Val Gly Ala Glu Ala Pro Gln His
 225 230 235 240
 Val Pro Pro Leu Arg Leu Leu Pro Ala Leu Val Trp Ala Ala Pro Gly
 245 250 255
 Val Ala Arg Gly Pro Val Val Ala Ser His Ala Leu Leu His Ala Pro
 260 265 270
 Pro Ala Gln Ala Ala Ala Pro Ser Pro Phe Trp Glu Gly His Ser Ala
 275 280 285

Ser Arg Gln His Glu Lys Leu Ser Arg Asn Ser Ser Thr Ser Glu Ser
 290 295 300

Ala Val Ser Ser Leu Ser Cys Pro Ala Arg Ala Trp Ala Ala Ala Ala
 305 310 315 320

Pro Cys Ala Ala

<210> 500
 <211> 23
 <212> PRT
 <213> Homo sapiens

<400> 500
 Glu Ala Val His Leu Pro Pro Val Leu Val Glu Gly Arg Gln Leu Leu
 1 5 10 15

Arg Val Arg Val Gln Gln Val
 20

<210> 501
 <211> 24
 <212> PRT
 <213> Homo sapiens

<400> 501
 Gly His Leu Glu Ala Ser Ala Glu Gly Leu Ala Arg Arg Gly Gly Gln
 1 5 10 15

Ala Gly Val Val Gly Val His Pro
 20

<210> 502
 <211> 28
 <212> PRT
 <213> Homo sapiens

<400> 502
 Gln Leu Glu Leu Ala Ala Glu Gly Gly Asp Gln Ala His Glu Gly Val
 1 5 10 15

Ala His Glu Glu Glu Leu Gly Val Leu Leu Glu Leu
 20 25

<210> 503
 <211> 27
 <212> PRT
 <213> Homo sapiens

<400> 503
 Gly Glu Leu Pro Val Ala Ala Pro Glu Leu Val Glu Gly Gln Val Arg

WO 99/31117

PCT/US98/27059

226

1

5

10

15

Ala Gly Val Val His Val Leu Ala Arg Asp Ala
20 25

<210> 504

<211> 25

<212> PRT

<213> Homo sapiens

<400> 504

Ala Val Gln Gln Ala Ser Ala Gln His Asp His His Ala Leu Pro Val
1 5 10 15

Gly Ala Gly His Leu Gly His Val Ala
20 25

<210> 505

<211> 25

<212> PRT

<213> Homo sapiens

<400> 505

His Asp Gln Val Ala Gln Leu Arg Val Gly Asp Val Val Glu Cys Ala
1 5 10 15

Leu Leu Gly Gly Glu Gly Gln Ala Gly
20 25

<210> 506

<211> 23

<212> PRT

<213> Homo sapiens

<400> 506

Ala Leu Val Trp Ala Ala Pro Gly Val Ala Arg Gly Pro Val Val Ala
1 5 10 15

Ser His Ala Leu Leu His Ala
20

<210> 507

<211> 28

<212> PRT

<213> Homo sapiens

<400> 507

Pro Pro Ala Gln Ala Ala Ala Pro Ser Pro Phe Trp Glu Gly His Ser
1 5 10 15

Ala Ser Arg Gln His Glu Lys Leu Ser Arg Asn Ser
20 25

SUBSTITUTE SHEET (RULE 26)

<210> 508
 <211> 314
 <212> PRT
 <213> Homo sapiens

<400> 508
 Ser Arg Val Thr Phe Pro Glu Arg Arg Arg Ser Ser Arg Leu Arg Arg
 1 5 10 15
 Gly Ser Met Glu Glu Ser Val Arg Gly Tyr Asp Trp Ser Pro Arg Asp
 20 25 30
 Ala Arg Arg Ser Pro Asp Gln Gly Arg Gln Gln Ala Glu Arg Arg Asn
 35 40 45
 Val Leu Arg Gly Phe Cys Ala Asn Ser Ser Leu Ala Phe Pro Thr Lys
 50 55 60
 Glu Arg Ala Phe Asp Asp Ile Pro Asn Ser Glu Leu Ser His Leu Ile
 65 70 75 80
 Val Asp Asp Arg His Gly Ala Ile Tyr Cys Tyr Val Pro Lys Val Ala
 85 90 95
 Cys Thr Asn Trp Lys Arg Val Met Ile Val Leu Ser Gly Ser Leu Leu
 100 105 110
 His Arg Gly Ala Pro Tyr Arg Asp Pro Leu Arg Ile Pro Arg Glu His
 115 120 125
 Val His Asn Ala Ser Ala His Leu Thr Phe Asn Lys Phe Trp Arg Arg
 130 135 140
 Tyr Gly Lys Leu Ser Arg His Leu Met Lys Val Lys Leu Lys Lys Tyr
 145 150 155 160
 Thr Lys Phe Leu Phe Val Arg Asp Pro Phe Val Arg Leu Ile Ser Ala
 165 170 175
 Phe Arg Ser Lys Phe Glu Leu Glu Asn Glu Glu Phe Tyr Arg Lys Phe
 180 185 190
 Ala Val Pro Met Leu Arg Val Tyr Ala Asn His Thr Ser Leu Pro Ala
 195 200 205
 Ser Ala Arg Glu Ala Phe Arg Ala Gly Leu Lys Val Ser Phe Ala Asn
 210 215 220
 Phe Ile Gln Tyr Leu Leu Asp Pro His Thr Glu Lys Leu Ala Pro Phe
 225 230 235 240
 Asn Glu His Trp Arg Gln Val Tyr Arg Leu Cys His Pro Cys Gln Ile
 245 250 255
 Asp Tyr Asp Ser Trp Gly Ser Trp Arg Leu Trp Thr Arg Thr Pro Arg
 260 265 270

Ser Cys Cys Ser Tyr Ser Arg Trp Thr Gly Ser Pro Leu Pro Pro Glu
 275 280 285

Leu Pro Glu Gln Asp Arg Gln Gln Leu Gly Gly Gly Leu Val Arg Gln
 290 295 300

Asp Pro Pro Gly Leu Glu Ala Ala Ala Val
 305 310

<210> 509
 <211> 26
 <212> PRT
 <213> Homo sapiens

<400> 509
 Arg Ser Pro Asp Gln Gly Arg Gln Gln Ala Glu Arg Arg Asn Val Leu
 1 5 10 15

Arg Gly Phe Cys Ala Asn Ser Ser Leu Ala
 20 25

<210> 510
 <211> 28
 <212> PRT
 <213> Homo sapiens

<400> 510
 Thr Lys Glu Arg Ala Phe Asp Asp Ile Pro Asn Ser Glu Leu Ser His
 1 5 10 15

Leu Ile Val Asp Asp Arg His Gly Ala Ile Tyr Cys
 20 25

<210> 511
 <211> 23
 <212> PRT
 <213> Homo sapiens

<400> 511
 Phe Asn Lys Phe Trp Arg Arg Tyr Gly Lys Leu Ser Arg His Leu Met
 1 5 10 15

Lys Val Lys Leu Lys Lys Tyr
 20

<210> 512
 <211> 24
 <212> PRT
 <213> Homo sapiens

<400> 512
 Phe Val Arg Leu Ile Ser Ala Phe Arg Ser Lys Phe Glu Leu Glu Asn

WO 99/31117

PCT/US98/27059

229

1

5

10

15

Glu Glu Phe Tyr Arg Lys Phe Ala
20

<210> 513

<211> 26

<212> PRT

<213> Homo sapiens

<400> 513

Thr Ser Leu Pro Ala Ser Ala Arg Glu Ala Phe Arg Ala Gly Leu Lys
1 5 10 15

Val Ser Phe Ala Asn Phe Ile Gln Tyr Leu
20 25

<210> 514

<211> 25

<212> PRT

<213> Homo sapiens

<400> 514

Ser Tyr Ser Arg Trp Thr Gly Ser Pro Leu Pro Pro Glu Leu Pro Glu
1 5 10 15

Gln Asp Arg Gln Gln Leu Gly Gly Gly
20 25

<210> 515

<211> 6

<212> PRT

<213> Homo sapiens

<400> 515

Ser Thr Gly Cys Ser Glu
1 5

<210> 516

<211> 146

<212> PRT

<213> Homo sapiens

<400> 516

Cys Leu Cys Leu Gly Cys Gly Leu Pro Glu Leu His Ser Tyr Leu Asp
1 5 10 15

Pro Gly Pro Tyr Leu Leu Val Tyr Pro Thr Leu Phe Trp Leu Cys Pro
20 25 30

Ser Ala Val Ser Pro Trp Ala Tyr Thr Cys Tyr Gln Leu Gly Leu Gly
35 40 45

SUBSTITUTE SHEET (RULE 26)

230

Pro Gln Trp Gly Ala Ala Ala Leu Ser Phe Thr Val Asp Ala Ala Ile
 50 55 60

Arg Val Trp Asp Val Ser Thr Glu Thr Cys Val Pro Leu Pro Trp Phe
 65 70 75 80

Arg Gly Gly Gly Val Thr Asn Cys Ser Gly Pro Gln Thr Ala Ala Lys
 85 90 95

Ser Trp Leu Pro Leu Leu Gln Leu Ser Phe Glu Ser Gly Arg Pro Arg
 100 105 110

Cys Gly Leu Val Arg Gly Gly Leu Leu Tyr Gln Gly Ala Val Arg Leu
 115 120 125

Ala Ala Gly Ala Gln Met Ala Ala Asp Cys Cys Ser Leu Tyr Trp Glu
 130 135 140

Ser His
 145

<210> 517
 <211> 26
 <212> PRT
 <213> Homo sapiens

<400> 517
 Tyr Pro Thr Leu Phe Trp Leu Cys Pro Ser Ala Val Ser Pro Trp Ala
 1 5 10 15

Tyr Thr Cys Tyr Gln Leu Gly Leu Gly Pro
 20 25

<210> 518
 <211> 25
 <212> PRT
 <213> Homo sapiens

<400> 518
 Asp Val Ser Thr Glu Thr Cys Val Pro Leu Pro Trp Phe Arg Gly Gly
 1 5 10 15

Gly Val Thr Asn Cys Ser Gly Pro Gln
 20 25

<210> 519
 <211> 22
 <212> PRT
 <213> Homo sapiens

<400> 519
 Leu Leu Tyr Gln Gly Ala Val Arg Leu Ala Ala Gly Ala Gln Met Ala
 1 5 10 15

231

Ala Asp Cys Cys Ser Leu
20

<210> 520
<211> 155
<212> PRT
<213> Homo sapiens

<400> 520
Asn Lys Arg Lys Thr Tyr Leu Phe Leu Glu Val Gly Met Trp Gly Val
1 5 10 15
Gly Gln Asn Arg Trp Trp Pro Trp Glu Arg Val Pro Arg Gly Arg Gly
20 25 30
Trp Gly Cys Leu Ser Lys Glu Gly Gln Val Met Asn Arg Ala Ser Thr
35 40 45
Pro Ser Arg Gly Phe Leu Gly Pro Pro Lys His Trp Ala Lys Thr Trp
50 55 60
Lys Leu Gly Ile Asp Lys Val Gln Arg Asp Val Gly Asn Ser Ala Cys
65 70 75 80
Gly Pro Ala His Thr Glu Gln Gly Pro Phe Val Glu Gly Arg Trp Lys
85 90 95
Val Met Ser Trp Gly Trp Ala Pro Gly Ser Pro Trp Ile Met Pro Gln
100 105 110
Gly Arg Ser Ser Asn Thr Gly Leu Phe Arg Val Arg Lys Arg Arg Met
115 120 125
Thr Gly Leu Pro Ser Cys Thr Leu Gly Phe Pro Phe Ile Ser Thr Ala
130 135 140
Arg Arg Ser Pro Leu Gly Ser Gln Thr Met Glu
145 150 155

<210> 521
<211> 26
<212> PRT
<213> Homo sapiens

<400> 521
Gly Val Gly Gln Asn Arg Trp Trp Pro Trp Glu Arg Val Pro Arg Gly
1 5 10 15
Arg Gly Trp Gly Cys Leu Ser Lys Glu Gly
20 25

<210> 522
<211> 26
<212> PRT

232

<213> Homo sapiens

<400> 522

Ala Lys Thr Trp Lys Leu Gly Ile Asp Lys Val Gln Arg Asp Val Gly
 1 5 10 15
 Asn Ser Ala Cys Gly Pro Ala His Thr Glu
 20 25

<210> 523

<211> 42

<212> PRT

<213> Homo sapiens

<400> 523

Trp Ala Pro Gly Ser Pro Trp Ile Met Pro Gln Gly Arg Ser Ser Asn
 1 5 10 15
 Thr Gly Leu Phe Arg Val Arg Lys Arg Arg Met Thr Gly Leu Pro Ser
 20 25 30
 Cys Thr Leu Gly Phe Pro Phe Ile Ser Thr
 35 40

<210> 524

<211> 17

<212> PRT

<213> Homo sapiens

<400> 524

Ser Ser Tyr Gln Cys Pro Lys Val Thr Phe Phe Lys Ser Ser Val Asp
 1 5 10 15

Thr

<210> 525

<211> 14

<212> PRT

<213> Homo sapiens

<400> 525

Tyr Ile Tyr Ser Tyr Leu Gly Phe Phe Asn Gln Ile Asn Lys
 1 5 10

<210> 526

<211> 6

<212> PRT

<213> Homo sapiens

<400> 526

Ala Arg Asp Leu Ile Leu
 1 5

<210> 527
<211> 43
<212> PRT
<213> Homo sapiens

<400> 527
Leu Thr Phe Tyr Leu Gln Phe Leu Ala Pro Lys Asp Lys Pro Ser Gly
1 5 10 15
Asp Thr Ala Ala Val Phe Glu Glu Gly Gly Asp Val Asp Asp Leu Val
20 25 30
Ser Thr Phe Asn Met His Leu Val Phe Cys Asp
35 40

<210> 528
<211> 25
<212> PRT
<213> Homo sapiens

<400> 528
Phe Leu Ala Pro Lys Asp Lys Pro Ser Gly Asp Thr Ala Ala Val Phe
1 5 10 15
Glu Glu Gly Gly Asp Val Asp Asp Leu
20 25

<210> 529
<211> 13
<212> PRT
<213> Homo sapiens

<400> 529
Ala Arg Ala Gly Ala Lys Ile Leu Phe Glu Gly Glu Phe
1 5 10

<210> 530
<211> 92
<212> PRT
<213> Homo sapiens

<400> 530
Asn Phe Glu Ile His Ser Ala Phe Pro Phe Met Leu Phe Val Ala Cys
1 5 10 15
Leu Leu His Ser Ser Cys Pro Arg Thr Ala Arg Phe Leu Ala Ser Pro
20 25 30
Leu Ser Glu Ser Asn Val Ile Phe Tyr Gln Asn Gln Tyr Gln Phe Pro
35 40 45
Cys Ile Leu Cys Phe Ile Glu Phe Ala Arg Leu Thr Ser Phe Lys His

234

50 55 60

Leu Ile His Ser Gln Ser His Leu Val Arg Leu Gln Tyr Glu Asp Phe
 65 70 75 80

Ser Val Ser Ser Glu Ala Trp Asp Thr Glu Leu Thr
 85 90

<210> 531
 <211> 26
 <212> PRT
 <213> Homo sapiens

<400> 531
 Phe Pro Phe Met Leu Phe Val Ala Cys Leu Leu His Ser Ser Cys Pro
 1 5 10 15

Arg Thr Ala Arg Phe Leu Ala Ser Pro Leu
 20 25

<210> 532
 <211> 26
 <212> PRT
 <213> Homo sapiens

<400> 532
 Asn Val Ile Phe Tyr Gln Asn Gln Tyr Gln Phe Pro Cys Ile Leu Cys
 1 5 10 15

Phe Ile Glu Phe Ala Arg Leu Thr Ser Phe
 20 25

<210> 533
 <211> 23
 <212> PRT
 <213> Homo sapiens

<400> 533
 Ser Gln Ser His Leu Val Arg Leu Gln Tyr Glu Asp Phe Ser Val Ser
 1 5 10 15

Ser Glu Ala Trp Asp Thr Glu
 20

<210> 534
 <211> 10
 <212> PRT
 <213> Homo sapiens

<400> 534
 Gln Lys Phe Leu Cys Ala Ser Asp Gly Asp
 1 5 10

<210> 535
 <211> 177
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (160)
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>
 <221> SITE
 <222> (162)
 <223> Xaa equals any of the naturally occurring L-amino acids

<400> 535
 Ala Glu Val Pro Leu Arg Val Arg Arg Arg His Gly Arg Pro His Gly
 1 5 10 15
 Pro Gly Gly Arg Gln Leu Ala Leu Gly Ile Pro Ala Leu Arg Ser Leu
 20 25 30
 Pro Gly Cys Val Pro Arg His His Gly Cys Ser Pro Gly Tyr Gly Cys
 35 40 45
 Leu His Arg Arg Ile Leu Cys Leu Pro Leu Ile Leu Leu Leu Val Tyr
 50 55 60
 Lys Gln Arg Gln Ala Ala Ser Asn Arg Arg Ala Gln Glu Leu Val Arg
 65 70 75 80
 Met Asp Ser Asn Ile Gln Gly Ile Glu Asn Pro Gly Phe Glu Ala Ser
 85 90 95
 Pro Pro Ala Gln Gly Ile Pro Glu Ala Lys Val Arg His Pro Leu Ser
 100 105 110
 Tyr Val Ala Gln Arg Gln Pro Ser Glu Ser Gly Arg His Leu Leu Ser
 115 120 125
 Glu Pro Ser Thr Pro Leu Ser Pro Pro Gly Pro Gly Asp Val Phe Phe
 130 135 140
 Pro Ser Leu Asp Pro Val Pro Asp Ser Pro Asn Phe Glu Val Ile Xaa
 145 150 155 160
 Pro Xaa Trp Gly Thr Val Gly Cys Cys Gly Trp Val Trp Gly Arg Cys
 165 170 175

Ile

<210> 536
 <211> 27
 <212> PRT

236

<213> Homo sapiens

<400> 536

Gly Pro Gly Gly Arg Gln Leu Ala Leu Gly Ile Pro Ala Leu Arg Ser
 1 5 10 15

Leu Pro Gly Cys Val Pro Arg His His Gly Cys
 20 25

<210> 537

<211> 25

<212> PRT

<213> Homo sapiens

<400> 537

Phe Glu Ala Ser Pro Pro Ala Gln Gly Ile Pro Glu Ala Lys Val Arg
 1 5 10 15

His Pro Leu Ser Tyr Val Ala Gln Arg
 20 25

<210> 538

<211> 88

<212> PRT

<213> Homo sapiens

<400> 538

Asp Met Ser Leu Gly Met Trp Gln His Gln Trp Asp Lys Met Asp Thr
 1 5 10 15

Gly Pro Pro Ser Gln Ala Pro Asp Thr Gly His Gly Gly Glu Thr Ser
 20 25 30

Pro Pro Trp His Ala Leu Gly Ser Pro Val Leu Pro Glu Ala Ala Leu
 35 40 45

Leu Ser Asp Phe Leu Phe Val Pro Gln Trp Leu Trp Gly Gln Ala Cys
 50 55 60

Leu Pro Thr Gly His Arg His Leu Pro Gln Leu Pro Pro Thr Ser Ser
 65 70 75 80

Phe Ser Glu Asp Leu Ser Thr Gly
 85

<210> 539

<211> 78

<212> PRT

<213> Homo sapiens

<400> 539

Pro Val Asp Arg Ser Ser Glu Lys Leu Leu Val Gly Gly Ser Trp Gly
 1 5 10 15

237

Arg Trp Arg Trp Pro Val Gly Arg Gln Ala Trp Pro Gln Ser His Cys
 20 25 30

Gly Thr Lys Arg Lys Ser Asp Arg Arg Ala Ala Ser Gly Lys Thr Gly
 35 40 45

Glu Pro Ser Ala Cys His Gly Gly Glu Val Ser Pro Pro Cys Pro Val
 50 55 60

Ser Gly Ala Trp Glu Gly Gly Pro Val Ser Ile Leu Ser His
 65 70 75

<210> 540

<211> 22

<212> PRT

<213> Homo sapiens

<400> 540

Pro Val Asp Arg Ser Ser Glu Lys Leu Leu Val Gly Gly Ser Trp Gly
 1 5 10 15

Arg Trp Arg Trp Pro Val
 20

<210> 541

<211> 25

<212> PRT

<213> Homo sapiens

<400> 541

Thr Lys Arg Lys Ser Asp Arg Arg Ala Ala Ser Gly Lys Thr Gly Glu
 1 5 10 15

Pro Ser Ala Cys His Gly Gly Glu Val
 20 25

<210> 542

<211> 46

<212> PRT

<213> Homo sapiens

<400> 542

Met Thr Ser Lys Phe Gly Glu Ser Gly Thr Gly Ser Arg Asp Gly Lys
 1 5 10 15

Lys Thr Ser Pro Gly Pro Gly Gly Asp Arg Gly Val Leu Gly Ser Glu
 20 25 30

Ser Arg Cys Arg Pro Asp Ser Glu Gly Cys Arg Trp Ala Thr
 35 40 45

<210> 543

<211> 20

<212> PRT

<213> Homo sapiens

<400> 543

Ser Pro Gly Pro Gly Gly Asp Arg Gly Val Leu Gly Ser Glu Ser Arg
 1 5 10 15

Cys Arg Pro Asp
 20

<210> 544

<211> 23

<212> PRT

<213> Homo sapiens

<400> 544

Pro Pro Ser Gln Ala Pro Asp Thr Gly His Gly Gly Glu Thr Ser Pro
 1 5 10 15

Pro Trp His Ala Leu Gly Ser
 20

<210> 545

<211> 15

<212> PRT

<213> Homo sapiens

<400> 545

His Glu Val Gln Pro Ser Tyr Leu Pro Ser Asn Ser Gly Leu Ile
 1 5 10 15

<210> 546

<211> 22

<212> PRT

<213> Homo sapiens

<400> 546

Leu Arg Ile Ser Val Leu Cys Arg Glu Thr Ala Cys Asn Trp Ser His
 1 5 10 15

His Pro Leu Asp Ser Asn
 20

<210> 547

<211> 32

<212> PRT

<213> Homo sapiens

<400> 547

Leu Thr Val Thr Val Arg Asn Pro Gly Ser Thr His Ala Ser Gly Arg
 1 5 10 15

Pro Arg Arg Arg Ser Gly Val Trp Ala Arg Arg Gly Leu Val Trp Gln

239

20

25

30

<210> 548
 <211> 38
 <212> PRT
 <213> Homo sapiens

<400> 548
 Thr Pro Cys Ser Ala Gln Phe Ser Val Leu Gly Pro Ser Gly Pro Ile
 1 5 10 15
 Leu Ala Met Val Gly Glu Asp Ala Asp Leu Pro Cys His Leu Phe Pro
 20 25 30
 Thr Met Ser Ala Glu Thr
 35

<210> 549
 <211> 60
 <212> PRT
 <213> Homo sapiens

<400> 549
 Met Glu Leu Lys Trp Val Ser Ser Ser Leu Arg Gln Val Val Asn Val
 1 5 10 15
 Tyr Ala Asp Gly Lys Glu Val Glu Asp Arg Gln Ser Ala Pro Tyr Arg
 20 25 30
 Gly Arg Thr Ser Ile Leu Arg Asp Gly Ile Thr Ala Gly Lys Ala Ala
 35 40 45
 Leu Arg Ile His Asn Val Thr Ala Ser Asp Ser Gly
 50 55 60

<210> 550
 <211> 26
 <212> PRT
 <213> Homo sapiens

<400> 550
 Leu Glu Val Lys Gly Tyr Glu Asp Gly Gly Ile His Leu Glu Cys Arg
 1 5 10 15
 Ser Thr Gly Trp Tyr Pro Gln Pro Gln Ile
 20 25

<210> 551
 <211> 80
 <212> PRT

<213> Homo sapiens

<400> 551

Met Ala Ser Ser Leu Ala Phe Leu Leu Leu Asn Phe His Val Ser Leu
 1 5 10 15
 Leu Leu Val Gln Leu Leu Thr Pro Cys Ser Ala Gln Phe Ser Val Leu
 20 25 30
 Gly Pro Ser Gly Pro Ile Leu Ala Met Val Gly Glu Asp Ala Asp Leu
 35 40 45
 Pro Cys His Leu Phe Pro Thr Met Ser Ala Glu Thr Met Glu Leu Lys
 50 55 60
 Trp Val Ser Ser Ser Leu Arg Gln Val Val Asn Val Tyr Ala Asp Gly
 65 70 75 80

<210> 552

<211> 103

<212> PRT

<213> Homo sapiens

<400> 552

Arg His Glu Leu Ser His Asn Arg Lys Asn Gly Glu Leu Leu Ile Asp
 1 5 10 15
 Arg Leu Tyr Ser Val Gly Ser Asp Ser Pro Met Gly Ile Pro Arg Asp
 20 25 30
 Ile Ile Phe Thr Asp Gly Phe Pro Tyr Trp Asn Pro Lys Val Lys Thr
 35 40 45
 Leu Lys Asp Arg His Phe Trp Gln Ser Ile Asp Glu Asn Gly Lys Phe
 50 55 60
 Pro Gly Phe Pro Ser Ala Gln Leu Ser Cys Leu Pro Pro Leu Gly Pro
 65 70 75 80
 Ala Ala His Ser Leu Leu Ser Ser Val Phe Cys Ala Trp Thr Leu Trp
 85 90 95
 Ala His Pro Gly His Gly Gly
 100

<210> 553

<211> 24

<212> PRT

<213> Homo sapiens

<400> 553

Leu Leu Ile Asp Arg Leu Tyr Ser Val Gly Ser Asp Ser Pro Met Gly

WO 99/31117

PCT/US98/27059

241

1

5

10

15

Ile Pro Arg Asp Ile Ile Phe Thr
20

<210> 554

<211> 25

<212> PRT

<213> Homo sapiens

<400> 554

Asn Pro Lys Val Lys Thr Leu Lys Asp Arg His Phe Trp Gln Ser Ile
1 5 10 15

Asp Glu Asn Gly Lys Phe Pro Gly Phe
20 25

<210> 555

<211> 24

<212> PRT

<213> Homo sapiens

<400> 555

Leu Gly Pro Ala Ala His Ser Leu Leu Ser Ser Val Phe Cys Ala Trp
1 5 10 15

Thr Leu Trp Ala His Pro Gly His
20

<210> 556

<211> 135

<212> PRT

<213> Homo sapiens

<400> 556

Arg Leu Gln His Trp Val Leu Ile Phe Thr Leu Glu Val Lys Gly Tyr
1 5 10 15

Glu Asp Gly Gly Ile His Leu Glu Cys Arg Ser Thr Gly Trp Tyr Pro
20 25 30

Gln Pro Gln Ile Gln Trp Ser Asn Ala Lys Gly Glu Asn Ile Pro Ala
35 40 45

Val Glu Ala Pro Val Val Ala Asp Gly Val Gly Leu Tyr Glu Val Ala
50 55 60

Ala Ser Val Ile Met Arg Gly Gly Ser Gly Glu Gly Val Ser Cys Ile
65 70 75 80

Ile Arg Asn Ser Leu Leu Gly Leu Glu Lys Thr Ala Ser Ile Ser Ile
85 90 95

Ala Asp Pro Ser Ser Gly Ala Pro Ser Pro Gly Ser Gln Pro Trp Gln

SUBSTITUTE SHEET (RULE 26)

242

100 105 110

Gly Pro Cys Leu Ser Cys Cys Cys Phe Ser Pro Glu Pro Val Thr Ser
 115 120 125

Cys Gly Asp Asn Arg Arg Lys
 130 135

<210> 557
 <211> 25
 <212> PRT
 <213> Homo sapiens

<400> 557
 Gly Gly Ile His Leu Glu Cys Arg Ser Thr Gly Trp Tyr Pro Gln Pro
 1 5 10 15

Gln Ile Gln Trp Ser Asn Ala Lys Gly
 20 25

<210> 558
 <211> 27
 <212> PRT
 <213> Homo sapiens

<400> 558
 Pro Gln Ile Gln Trp Ser Asn Ala Lys Gly Glu Asn Ile Pro Ala Val
 1 5 10 15

Glu Ala Pro Val Val Ala Asp Gly Val Gly Leu
 20 25

<210> 559
 <211> 27
 <212> PRT
 <213> Homo sapiens

<400> 559
 Asn Ile Pro Ala Val Glu Ala Pro Val Val Ala Asp Gly Val Gly Leu
 1 5 10 15

Tyr Glu Val Ala Ala Ser Val Ile Met Arg Gly
 20 25

<210> 560
 <211> 27
 <212> PRT
 <213> Homo sapiens

<400> 560
 Ser Gly Ala Pro Ser Pro Gly Ser Gln Pro Trp Gln Gly Pro Cys Leu
 1 5 10 15

243

Ser Cys Cys Cys Phe Ser Pro Glu Pro Val Thr
 20 25

<210> 561
 <211> 131
 <212> PRT
 <213> Homo sapiens

<400> 561
 Ser Ser Ser Ile Cys Asp His Glu Arg Arg Leu Arg Gly Gly Cys Ile
 1 5 10 15
 Leu His His Gln Lys Phe Pro Pro Arg Pro Gly Lys Asp Ser Gln His
 20 25 30
 Phe His Arg Arg Pro Phe Phe Arg Ser Ala Gln Pro Trp Ile Ala Ala
 35 40 45
 Leu Ala Gly Thr Leu Pro Ile Leu Leu Leu Leu Leu Ala Gly Ala Ser
 50 55 60
 Tyr Phe Leu Trp Arg Gln Gln Lys Glu Ile Thr Ala Leu Ser Ser Glu
 65 70 75 80
 Ile Glu Ser Glu Gln Glu Met Lys Glu Met Gly Tyr Ala Ala Thr Glu
 85 90 95
 Arg Glu Ile Ser Leu Arg Glu Ser Leu Gln Glu Glu Leu Lys Arg Lys
 100 105 110
 Lys Ile Gln Tyr Leu Thr Arg Gly Glu Glu Ser Ser Ser Asp Thr Asn
 115 120 125
 Lys Ser Ala
 130

<210> 562
 <211> 28
 <212> PRT
 <213> Homo sapiens

<400> 562
 Lys Asp Ser Gln His Phe His Arg Arg Pro Phe Phe Arg Ser Ala Gln
 1 5 10 15
 Pro Trp Ile Ala Ala Leu Ala Gly Thr Leu Pro Ile
 20 25

<210> 563
 <211> 28
 <212> PRT
 <213> Homo sapiens

<400> 563

WO 99/31117

PCT/US98/27059

244

Glu Ile Glu Ser Glu Gln Glu Met Lys Glu Met Gly Tyr Ala Ala Thr
1 5 10 15

Glu Arg Glu Ile Ser Leu Arg Glu Ser Leu Gln Glu
20 25

<210> 564

<211> 33

<212> PRT

<213> Homo sapiens

<400> 564

Val Asn Asn Met Ile Ala Phe Tyr Ser Ala Arg Asp Ser Tyr Val Tyr
1 5 10 15

Pro His Phe Ser Gly Glu Glu Met Leu Gln Met Arg Leu His Leu Val
20 25 30

Lys

<210> 565

<211> 38

<212> PRT

<213> Homo sapiens

<400> 565

Thr Pro Cys Ser Ala Gln Phe Ser Val Leu Gly Pro Ser Gly Pro Ile
1 5 10 15

Leu Ala Met Val Gly Glu Asp Ala Asp Leu Pro Cys His Leu Phe Pro
20 25 30

Thr Met Ser Ala Glu Thr
35

<210> 566

<211> 23

<212> PRT

<213> Homo sapiens

<400> 566

Lys Trp Val Ser Ser Ser Leu Arg Gln Val Val Asn Val Tyr Ala Asp
1 5 10 15

Gly Lys Glu Val Glu Asp Arg
20

<210> 567

<211> 25

<212> PRT

<213> Homo sapiens

SUBSTITUTE SHEET (RULE 26)

245

<400> 567

Arg Thr Ser Ile Leu Arg Asp Gly Ile Thr Ala Gly Lys Ala Ala Leu
1 5 10 15

Arg Ile His Asn Val Thr Ala Ser Asp
20 25

<210> 568

<211> 23

<212> PRT

<213> Homo sapiens

<400> 568

Cys Tyr Phe Gln Asp Gly Asp Phe Tyr Glu Lys Ala Leu Val Glu Leu
1 5 10 15

Lys Val Ala Ala Leu Gly Ser
20

<210> 569

<211> 23

<212> PRT

<213> Homo sapiens

<400> 569

Gly Tyr Glu Asp Gly Gly Ile His Leu Glu Cys Arg Ser Thr Gly Trp
1 5 10 15

Tyr Pro Gln Pro Gln Ile Gln
20

<210> 570

<211> 23

<212> PRT

<213> Homo sapiens

<400> 570

Asn Ile Pro Ala Val Glu Ala Pro Val Val Ala Asp Gly Val Gly Leu
1 5 10 15

Tyr Glu Val Ala Ala Ser Val
20

<210> 571

<211> 21

<212> PRT

<213> Homo sapiens

<400> 571

Gln Gln Lys Glu Ile Thr Ala Leu Ser Ser Glu Ile Glu Ser Glu Gln
1 5 10 15

Glu Met Lys Glu Met

20

<210> 572
<211> 24
<212> PRT
<213> Homo sapiens

<400> 572
Leu Arg Glu Ser Leu Gln Glu Glu Leu Lys Arg Lys Lys Ile Gln Tyr
1 5 10 15

Leu Thr Arg Gly Glu Glu Ser Ser
20

<210> 573
<211> 13
<212> PRT
<213> Homo sapiens

<400> 573
Gly Glu Glu Met Leu Gln Met Arg Leu His Leu Val Lys
1 5 10

<210> 574
<211> 40
<212> PRT
<213> Homo sapiens

<400> 574
Ser Ala Gln Phe Ser Val Leu Gly Pro Ser Gly Pro Ile Leu Ala Met
1 5 10 15

Val Gly Glu Asp Ala Asp Leu Pro Cys His Leu Phe Pro Thr Met Ser
20 25 30

Ala Glu Thr Met Glu Leu Lys Trp
35 40

<210> 575
<211> 12
<212> PRT
<213> Homo sapiens

<400> 575
Pro Gln Gly Gly Leu Thr Leu Pro Ser Val Trp Gly
1 5 10

<210> 576
<211> 106
<212> PRT
<213> Homo sapiens

247

<400> 576

Gly Gly Pro Cys His Leu Trp Leu Leu Gly Pro Arg Arg Thr Gln Leu
 1 5 10 15

Pro Gly Arg Arg Ala Ser Leu Pro Phe Arg Ser Gln Gly Glu Leu Thr
 20 25 30

Gln Ala Phe Leu Leu Gly Leu Trp Lys His Gln Met Pro Ala Leu Thr
 35 40 45

Gln Glu Gln Gln Val Arg Ala Glu Arg Arg Arg Glu Ala Val Arg Met
 50 55 60

Glu Ile Pro Gly Leu Phe Phe Ala Ser Leu Ala Asn Trp Gly Leu Leu
 65 70 75 80

Tyr Arg Thr Ser Gln Asp Phe Ile Ser Pro Tyr Leu Cys Ala Ala Pro
 85 90 95

Ser Thr Pro His Pro Pro Leu Gly Gly Pro
 100 105

<210> 577

<211> 23

<212> PRT

<213> Homo sapiens

<400> 577

Gly Pro Arg Arg Thr Gln Leu Pro Gly Arg Arg Ala Ser Leu Pro Phe
 1 5 10 15

Arg Ser Gln Gly Glu Leu Thr
 20

<210> 578

<211> 24

<212> PRT

<213> Homo sapiens

<400> 578

Gln Met Pro Ala Leu Thr Gln Glu Gln Gln Val Arg Ala Glu Arg Arg
 1 5 10 15

Arg Glu Ala Val Arg Met Glu Ile
 20

<210> 579

<211> 25

<212> PRT

<213> Homo sapiens

<400> 579

Ala Asn Trp Gly Leu Leu Tyr Arg Thr Ser Gln Asp Phe Ile Ser Pro
 1 5 10 15

Tyr Leu Cys Ala Ala Pro Ser Thr Pro
20 25

<210> 580
<211> 34
<212> PRT
<213> Homo sapiens

<400> 580
Leu Ser Phe Lys Asp Lys Ser Thr Tyr Ile Glu Ser Ser Thr Lys Val
1 5 10 15

Tyr Asp Asp Met Ala Phe Arg Tyr Leu Ser Trp Ile Leu Phe Pro Leu
20 25 30

Leu Gly

<210> 581
<211> 31
<212> PRT
<213> Homo sapiens

<400> 581
Leu Leu Thr Phe Gly Phe Ile Thr Met Thr Pro Gln Leu Phe Ile Asn
1 5 10 15

Tyr Lys Leu Lys Ser Val Ala His Leu Pro Trp Arg Met Leu Thr
20 25 30

<210> 582
<211> 30
<212> PRT
<213> Homo sapiens

<400> 582
Thr Tyr Lys Ala Leu Asn Thr Phe Ile Asp Asp Leu Phe Ala Phe Val
1 5 10 15

Ile Lys Met Pro Val Met Tyr Arg Ile Gly Cys Leu Arg Asp
20 25 30

<210> 583
<211> 30
<212> PRT
<213> Homo sapiens

<400> 583
Asp Val Val Phe Phe Ile Tyr Leu Tyr Gln Arg Trp Ile Tyr Arg Val
1 5 10 15

Asp Pro Thr Arg Val Asn Glu Phe Gly Met Ser Gly Glu Asp

249

20

25

30

<210> 584
 <211> 44
 <212> PRT
 <213> Homo sapiens

<400> 584
 Val Ala Gly Ile Phe Pro Arg Leu Ser Phe Lys Asp Lys Ser Thr Tyr
 1 5 10 15
 Ile Glu Ser Ser Thr Lys Val Tyr Asp Asp Met Ala Phe Arg Tyr Leu
 20 25 30
 Ser Trp Ile Leu Phe Pro Leu Leu Gly Cys Tyr Ala
 35 40

<210> 585
 <211> 19
 <212> PRT
 <213> Homo sapiens

<400> 585
 Trp Ala Ala Met Pro Ser Thr Val Phe Cys Thr Trp Ser Thr Arg Ala
 1 5 10 15
 Gly Thr Pro

<210> 586
 <211> 28
 <212> PRT
 <213> Homo sapiens

<400> 586
 Pro Trp Val Ala Gly Ile Phe Pro Arg Leu Ser Phe Lys Asp Lys Ser
 1 5 10 15
 Thr Tyr Ile Glu Ser Ser Thr Lys Val Tyr Asp Asp
 20 25

<210> 587
 <211> 88
 <212> PRT
 <213> Homo sapiens

<400> 587
 Ala Gly Glu Asp Ser Cys His Pro Val Leu Ser Val Gln Pro Asp Val
 1 5 10 15
 His Asp Leu Gly Trp Gln Glu Ser Ser Pro Ala Tyr Pro Ser Arg Thr
 20 25 30

250

Ser Pro Arg Ile Ser Ser Pro Arg Pro Lys Cys Met Met Ile Trp His
 35 40 45

Ser Gly Thr Cys Pro Gly Ser Ser Ser Arg Ser Trp Ala Ala Met Pro
 50 55 60

Ser Thr Val Phe Cys Thr Trp Ser Thr Arg Ala Gly Thr Pro Gly Cys
 65 70 75 80

Ser Ala Cys Ser Thr Ala Ser Cys
 85

<210> 588

<211> 30

<212> PRT

<213> Homo sapiens

<400> 588

Leu Ser Val Gln Pro Asp Val His Asp Leu Gly Trp Gln Glu Ser Ser
 1 5 10 15

Pro Ala Tyr Pro Ser Arg Thr Ser Pro Arg Ile Ser Ser Pro
 20 25 30

<210> 589

<211> 25

<212> PRT

<213> Homo sapiens

<400> 589

Gly Ser Ser Ser Arg Ser Trp Ala Ala Met Pro Ser Thr Val Phe Cys
 1 5 10 15

Thr Trp Ser Thr Arg Ala Gly Thr Pro
 20 25

<210> 590

<211> 22

<212> PRT

<213> Homo sapiens

<400> 590

Cys Tyr Ala Val Tyr Ser Leu Leu Tyr Leu Glu His Lys Gly Trp Tyr
 1 5 10 15

Ser Trp Val Leu Ser Met
 20

<210> 591

<211> 12

<212> PRT

<213> Homo sapiens

251

<400> 591

Leu Gly Glu Phe Leu Ser Ser Gln Cys Phe Leu Pro
 1 5 10

<210> 592

<211> 20

<212> PRT

<213> Homo sapiens

<400> 592

Arg Ser Arg Arg Asn Arg Val Ala Met Gly Met Trp Ala Ser Leu Asp
 1 5 10 15

Ala Leu Trp Glu
 20

<210> 593

<211> 92

<212> PRT

<213> Homo sapiens

<400> 593

Pro Arg Val Arg Cys Gln Gln Arg Ala Glu Gly Gly Met Gly Ala Gly
 1 5 10 15

Ile Gly Val Gly Pro Ser Glu Arg Thr Asp Ile Ala Val Thr Pro Arg
 20 25 30

Gly Arg Ser Glu Gly Ala Ser Val Gly Val Ala Pro Val His Ala Glu
 35 40 45

Gly Ala Gly Gly Thr Gly Trp Pro Trp Gly Cys Gly His Arg Trp Thr
 50 55 60

Leu Cys Gly Arg Cys Arg Pro Arg Ser Val Ser Ser Gly Pro Cys Cys
 65 70 75 80

Ser Phe Pro Gly Gln Cys Ile Phe Gly Arg Pro Ser
 85 90

<210> 594

<211> 24

<212> PRT

<213> Homo sapiens

<400> 594

Gly Gly Met Gly Ala Gly Ile Gly Val Gly Pro Ser Glu Arg Thr Asp
 1 5 10 15

Ile Ala Val Thr Pro Arg Gly Arg
 20

<210> 595

252

<211> 26
 <212> PRT
 <213> Homo sapiens

<400> 595
 Gly Cys Gly His Arg Trp Thr Leu Cys Gly Arg Cys Arg Pro Arg Ser
 1 5 10 15
 Val Ser Ser Gly Pro Cys Cys Ser Phe Pro
 20 25

<210> 596
 <211> 24
 <212> PRT
 <213> Homo sapiens

<400> 596
 Lys Lys His Gly Phe Asn Gln Gln Thr Leu Gly Phe Phe Thr Trp Lys
 1 5 10 15
 Tyr Asn Lys Asn Lys Asn Leu Val
 20

<210> 597
 <211> 21
 <212> PRT
 <213> Homo sapiens

<400> 597
 Pro Lys Leu Leu Pro Cys Ser Pro Ala Glu Gly His Thr Ser Leu Gly
 1 5 10 15
 Pro Leu Leu Pro Phe
 20

<210> 598
 <211> 70
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (6)
 <223> Xaa equals any of the naturally occurring L-amino acids

<400> 598
 Ala Ser Leu Glu Leu Xaa Pro Ser Lys Ser Gln Leu Ser Thr Glu Trp
 1 5 10 15
 Gly Phe Thr Trp Ile Val Gly Leu Gly Met Ser Pro Ser Thr Ala Leu
 20 25 30
 Trp Thr Glu Cys Thr Cys Thr Pro Phe Leu Val Leu Leu Ser His Ala
 35 40 45

Ser Gly His Phe Phe Trp Leu Ser Pro Leu Ala Ser Leu Val Ile Pro
 50 55 60

Pro Val Thr Asp Arg Lys
 65 70

<210> 599
 <211> 32
 <212> PRT
 <213> Homo sapiens

<400> 599
 Trp Gly Phe Thr Trp Ile Val Gly Leu Gly Met Ser Pro Ser Thr Ala
 1 5 10 15

Leu Trp Thr Glu Cys Thr Cys Thr Pro Phe Leu Val Leu Leu Ser His
 20 25 30

<210> 600
 <211> 106
 <212> PRT
 <213> Homo sapiens

<400> 600
 Val Ala Val Gly Val Cys Arg Glu Asp Val Met Gly Ile Thr Asp Arg
 1 5 10 15

Ser Lys Met Ser Pro Asp Val Gly Ile Trp Ala Ile Tyr Trp Ser Ala
 20 25 30

Ala Gly Tyr Trp Pro Leu Ile Gly Phe Pro Gly Thr Pro Thr Gln Gln
 35 40 45

Glu Pro Ala Leu His Arg Val Gly Val Tyr Leu Asp Arg Gly Thr Gly
 50 55 60

Asn Val Ser Phe Tyr Ser Ala Val Asp Gly Val His Leu His Thr Phe
 65 70 75 80

Ser Cys Ser Ser Val Ser Arg Leu Arg Pro Phe Phe Leu Val Glu Ser
 85 90 95

Ile Ser Ile Phe Ser His Ser Thr Ser Asp
 100 105

<210> 601
 <211> 27
 <212> PRT
 <213> Homo sapiens

254

<400> 601

Ile Thr Asp Arg Ser Lys Met Ser Pro Asp Val Gly Ile Trp Ala Ile
1 5 10 15

Tyr Trp Ser Ala Ala Gly Tyr Trp Pro Leu Ile
20 25

<210> 602

<211> 30

<212> PRT

<213> Homo sapiens

<400> 602

Arg Gly Thr Gly Asn Val Ser Phe Tyr Ser Ala Val Asp Gly Val His
1 5 10 15

Leu His Thr Phe Ser Cys Ser Ser Val Ser Arg Leu Arg Pro
20 25 30

<210> 603

<211> 11

<212> PRT

<213> Homo sapiens

<400> 603

Gly Thr Arg Gly Leu Gln Asn His Arg Thr Glu
1 5 10

<210> 604

<211> 6

<212> PRT

<213> Homo sapiens

<400> 604

Glu Leu Ser Gly Leu Gly
1 5

<210> 605

<211> 6

<212> PRT

<213> Homo sapiens

<400> 605

Met Asp Asp Ile Lys Ile
1 5

<210> 606

<211> 57

<212> PRT

<213> Homo sapiens

<400> 606

WO 99/31117

PCT/US98/27059

255

Asn Phe Cys Val Ser Lys Asn Thr Phe Asn Arg Val Lys Arg Pro Ile
1 5 10 15

Lys Trp Val Lys Ile Phe Ala Asn Asp Ile Ser Cys Lys Arg Leu Ile
20 25 30

Ser Arg Ile His Lys Glu Ile Leu Pro Phe Asn Asn Lys Lys Gln Pro
35 40 45

Asp Phe Lys Val Lys Lys Ser Arg Lys
50 55

<210> 607

<211> 30

<212> PRT

<213> Homo sapiens

<400> 607

Phe Asn Arg Val Lys Arg Pro Ile Lys Trp Val Lys Ile Phe Ala Asn
1 5 10 15

Asp Ile Ser Cys Lys Arg Leu Ile Ser Arg Ile His Lys Glu
20 25 30

<210> 608

<211> 15

<212> PRT

<213> Homo sapiens

<400> 608

Glu Thr Gln Met Ala Asn Lys Tyr Met Lys Arg Cys Ser Thr Leu
1 5 10 15

<210> 609

<211> 59

<212> PRT

<213> Homo sapiens

<400> 609

Val Ile Arg Glu Leu Gln Val Lys Ala Thr Arg Arg Cys His Tyr Thr
1 5 10 15

Pro Ile Lys Trp Ser Lys Ser Lys Thr Leu Ile Ser Ser Asn Ala Asp
20 25 30

Glu Tyr Val Glu Pro Thr Arg Thr Leu Ile His Cys Trp Trp Lys Cys
35 40 45

Lys Ile Val Gln Pro Leu Cys Lys Thr Ala Trp
50 55

<210> 610

<211> 22

SUBSTITUTE SHEET (RULE 26)

<212> PRT

<213> Homo sapiens

<400> 610

Ala Thr Arg Arg Cys His Tyr Thr Pro Ile Lys Trp Ser Lys Ser Lys
 1 5 10 15

Thr Leu Ile Ser Ser Asn
 20

<210> 611

<211> 64

<212> PRT

<213> Homo sapiens

<400> 611

Glu Leu Ser Gly Leu Val Ile Ile Thr Ala Trp Ile Ile Leu Cys His
 1 5 10 15

Ser Ser Ser Lys Asn Pro Val Gly Gly Arg Ile Gln Leu Ala Ile Ala
 20 25 30

Ile Val Ile Thr Leu Phe Pro Phe Ile Ser Trp Val Tyr Ile Tyr Ile
 35 40 45

Asn Lys Glu Met Arg Ser Ser Trp Pro Thr His Cys Lys Thr Val Ile
 50 55 60

<210> 612

<211> 57

<212> PRT

<213> Homo sapiens

<400> 612

Gln Cys Pro Gln Gly Thr Glu Thr Glu Ala Gly Val Ser Val Pro Pro
 1 5 10 15

Arg Lys Glu Gly Gly Gly Pro Tyr Val Ala Gly Leu Thr Ala Pro His
 20 25 30

Val Ala Gly Leu Thr Ala Pro Arg Arg Val Leu Arg Ala Met Ala Pro
 35 40 45

Ala Leu Trp Arg Ala Cys Asn Gly Leu
 50 55

<210> 613

<211> 32

<212> PRT

<213> Homo sapiens

257

<400> 613

His Ser Ser Ser Lys Asn Pro Val Gly Gly Arg Ile Gln Leu Ala Ile
1 5 10 15

Ala Ile Val Ile Thr Leu Phe Pro Phe Ile Ser Trp Val Tyr Ile Tyr
20 25 30

<210> 614

<211> 32

<212> PRT

<213> Homo sapiens

<400> 614

Arg Lys Glu Gly Gly Gly Pro Tyr Val Ala Gly Leu Thr Ala Pro His
1 5 10 15

Val Ala Gly Leu Thr Ala Pro Arg Arg Val Leu Arg Ala Met Ala Pro
20 25 30

<210> 615

<211> 32

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (9)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 615

Pro Gly Arg Pro Thr Arg Pro Ala Xaa Ala Gly Leu Ser Ser Gly Gly
1 5 10 15

Ala Ala Gln Glu Ala Pro Gln Ala Asp Pro Arg Pro Trp Leu Ala Arg
20 25 30

<210> 616

<211> 51

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (3)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (29)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 616

His Tyr Xaa Ser Thr Pro Gly Arg Val Pro Val Arg Gln Phe Ala Ala
 1 5 10 15

Ala Ser Thr Ser Gly Gly Pro Trp Val Pro Gly Gly Xaa Leu Glu Ala
 20 25 30

Pro Phe Gln Val Ala Pro Ser Leu Ser His Ser Thr Pro Val Phe Pro
 35 40 45

Gly Leu Ile
 50

<210> 617

<211> 22

<212> PRT

<213> Homo sapiens

<400> 617

Ala Arg Gly Lys Tyr Glu Ser Ala Gln Pro Gly Gly Thr Gln Pro Glu
 1 5 10 15

Pro Gly Leu Gly Ala Arg
 20

<210> 618

<211> 24

<212> PRT

<213> Homo sapiens

<400> 618

Ser Cys Gly Ser Ser Arg Arg Ser Ala Lys Arg Ser Leu Thr Leu Lys
 1 5 10 15

Leu Ile Asp Phe Ser His Arg Ile
 20

<210> 619

<211> 52

<212> PRT

<213> Homo sapiens

<400> 619

His Tyr Phe Leu Arg Thr Val Ser Gly Leu Ser Val Val Pro Val Ser
 1 5 10 15

Leu Arg Cys Cys Met Cys Pro Pro Pro Cys Thr Gly Pro Ala Pro Ala
 20 25 30

Thr Ala His Ser Pro Phe Asp Pro Pro Ala Leu Pro Ile Gln Phe Glu
35 40 45

Tyr Gln Gln Ala
50

<210> 620
<211> 45
<212> PRT
<213> Homo sapiens

<400> 620
Gln Leu Glu Ala Glu Ile Glu Asn Leu Ser Trp Lys Val Glu Arg Ala
1 5 10 15

Asp Ser Tyr Asp Arg Gly Asp Leu Glu Asn Gln Met His Ile Ala Glu
20 25 30

Gln Arg Arg Arg Thr Leu Leu Lys Asp Phe His Asp Thr
35 40 45

<210> 621
<211> 24
<212> PRT
<213> Homo sapiens

<400> 621
Val Pro Val Ser Leu Arg Cys Cys Met Cys Pro Pro Pro Cys Thr Gly
1 5 10 15

Pro Ala Pro Ala Thr Ala His Ser
20

<210> 622
<211> 25
<212> PRT
<213> Homo sapiens

<400> 622
Ser Trp Lys Val Glu Arg Ala Asp Ser Tyr Asp Arg Gly Asp Leu Glu
1 5 10 15

Asn Gln Met His Ile Ala Glu Gln Arg
20 25

<210> 623
<211> 227
<212> PRT
<213> Homo sapiens

<220>
<221> SITE

<222> (53)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 623

His Glu Ala Trp Leu Arg Ser Ala Gly Thr Arg Glu Pro Pro Arg Glu
 1 5 10 15

Gln Arg Thr Arg Arg Arg Gln Thr Ala Gln Leu Ala Leu Gln Val Pro
 20 25 30

Ala Pro Ser Arg Thr Pro Pro Met Ala Thr Asp Val Phe Asn Ser Lys
 35 40 45

Asn Leu Ala Val Xaa Ala Gln Lys Lys Ile Leu Gly Lys Met Val Ser
 50 55 60

Lys Ser Ile Ala Thr Thr Leu Ile Asp Asp Thr Ser Ser Glu Val Leu
 65 70 75 80

Asp Glu Leu Tyr Arg Val Thr Arg Glu Tyr Thr Gln Asn Lys Lys Glu
 85 90 95

Ala Glu Lys Ile Ile Lys Asn Leu Ile Lys Thr Val Ile Lys Leu Ala
 100 105 110

Ile Leu Tyr Arg Asn Asn Gln Phe Asn Gln Asp Glu Leu Ala Leu Met
 115 120 125

Glu Lys Phe Lys Lys Lys Val His Gln Leu Ala Met Thr Val Val Ser
 130 135 140

Phe His Gln Val Asp Tyr Thr Phe Asp Arg Asn Val Leu Ser Arg Leu
 145 150 155 160

Leu Asn Glu Cys Arg Glu Met Leu His Gln Ile Ile Gln Arg His Leu
 165 170 175

Thr Ala Lys Ser His Gly Arg Val Asn Asn Val Phe Asp His Phe Ser
 180 185 190

Asp Cys Glu Phe Leu Ala Ala Leu Tyr Asn Pro Phe Gly Asn Phe Lys
 195 200 205

Pro His Leu Gln Lys Leu Cys Asp Gly Ile Asn Lys Met Leu Asp Glu
 210 215 220

Glu Asn Ile
 225

<210> 624

<211> 52

<212> PRT

<213> Homo sapiens

<400> 624

His Glu Ala Trp Leu Arg Ser Ala Gly Thr Arg Glu Pro Pro Arg Glu

WO 99/31117

PCT/US98/27059

261

1 5 10 15

Gln Arg Thr Arg Arg Arg Gln Thr Ala Gln Leu Ala Leu Gln Val Pro
20 25 30

Ala Pro Ser Arg Thr Pro Pro Met Ala Thr Asp Val Phe Asn Ser Lys
35 40 45

Asn Leu Ala Val
50

<210> 625

<211> 49

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (1)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 625

Xaa Ala Gln Lys Lys Ile Leu Gly Lys Met Val Ser Lys Ser Ile Ala
1 5 10 15

Thr Thr Leu Ile Asp Asp Thr Ser Ser Glu Val Leu Asp Glu Leu Tyr
20 25 30

Arg Val Thr Arg Glu Tyr Thr Gln Asn Lys Lys Glu Ala Glu Lys Ile
35 40 45

Ile

<210> 626

<211> 51

<212> PRT

<213> Homo sapiens

<400> 626

Lys Asn Leu Ile Lys Thr Val Ile Lys Leu Ala Ile Leu Tyr Arg Asn
1 5 10 15

Asn Gln Phe Asn Gln Asp Glu Leu Ala Leu Met Glu Lys Phe Lys Lys
20 25 30

Lys Val His Gln Leu Ala Met Thr Val Val Ser Phe His Gln Val Asp
35 40 45

Tyr Thr Phe
50

<210> 627

<211> 52

SUBSTITUTE SHEET (RULE 26)

<212> PRT

<213> Homo sapiens

<400> 627

Asp Arg Asn Val Leu Ser Arg Leu Leu Asn Glu Cys Arg Glu Met Leu
1 5 10 15

His Gln Ile Ile Gln Arg His Leu Thr Ala Lys Ser His Gly Arg Val
20 25 30

Asn Asn Val Phe Asp His Phe Ser Asp Cys Glu Phe Leu Ala Ala Leu
35 40 45

Tyr Asn Pro Phe
50

<210> 628

<211> 23

<212> PRT

<213> Homo sapiens

<400> 628

Gly Asn Phe Lys Pro His Leu Gln Lys Leu Cys Asp Gly Ile Asn Lys
1 5 10 15

Met Leu Asp Glu Glu Asn Ile
20

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US98/27059

A. CLASSIFICATION OF SUBJECT MATTER IPC(6) :C07H 21/00; C12N 1/15, 1/21, 5/10, 15/11, 15/63 US CL :435/91.41, 320.1, 325, 252.3, 254.11; 536/23.1, 23.5, 24.31 According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) U.S. : 435/91.41, 320.1, 325, 252.3, 254.11; 536/23.1, 23.5, 24.31 Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) GENBANK, EMBL, SWISS-PROT, SPTREMBL, PIR, searched: SEQ ID NO: 11-20 & 125-134		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	Database GenBank, US National Library of Medicine, (Bethesda, MD, USA), No. AA133381, HILLIER et al. 'WashU-Merck EST Project', complete record, 27 November 1996.	1, 7-10
X	Database GenBank, US National Library of Medicine, (Bethesda, MD, USA), No. T12400, LIEW et al. 'A catalogue of genes in the cardiovascular system as identified by expressed sequence tags', complete record, 27 November 1996.	1, 7-10
X	Database GenBank, US National Library of Medicine, (Bethesda, MD, USA), No. AA496982, HILLIER et al. 'WashU-Merck EST Project 1997', complete record, 12 August 1997.	1, 7-10
<input checked="" type="checkbox"/> Further documents are listed in the continuation of Box C. <input type="checkbox"/> See patent family annex.		
* Special categories of cited documents: *A* document defining the general state of the art which is not considered to be of particular relevance *B* earlier document published on or after the international filing date *L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) *O* document referring to an oral disclosure, use, exhibition or other means *P* document published prior to the international filing date but later than the priority date claimed *T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention *X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone *Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art *A* document member of the same patent family		
Date of the actual completion of the international search 02 MARCH 1999		Date of mailing of the international search report 23 MAR 1999
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230		Authorized officer SCOTT D. PRIEBE Telephone No. (703) 308-0196

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US98/27059

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	Database GenBank, US National Library of Medicine, (Bethesda, MD, USA), No. U14626, D'ALESSIO et al. 'Cloning vector pSVSport1', complete record, 24 May 1995.	1, 7-10
X	Database GenBank, US National Library of Medicine, (Bethesda, MD, USA), No. AJ000730, SPERANDEO et al. 'The full cDNA for the human cationic amino acid transporter 3 (HCAT3)', complete record, 02 December 1997.	1
X	Database GenBank, US National Library of Medicine, (Bethesda, MD, USA), No. R31044, HILLIER et al. 'The WashU-Merck EST Project', complete record, 28 April 1995.	1, 7-10
X	Database GenBank, US National Library of Medicine, (Bethesda, MD, USA), No. AA446873, HILLIER et al. 'WashU-Merck EST Project 1997', complete record, 03 June 1997.	1, 7-10
X	Database GenBank, US National Library of Medicine, (Bethesda, MD, USA), No. AA135715, HILLIER et al. 'WashU-Merck EST Project', complete record, 14 May 1997.	1, 7-10
X	Database GenBank, US National Library of Medicine, (Bethesda, MD, USA), No. AA194015, HILLIER et al. 'WashU-Merck EST Project', complete record, 19 May 1997.	1, 7-10
X	Database GenBank, US National Library of Medicine, (Bethesda, MD, USA), No. R72850, HILLIER et al. 'The WashU-Merck EST Project', complete record, 02 June 1995.	1, 7-10
X	Database GenBank, US National Library of Medicine, (Bethesda, MD, USA), No. T60940, HILLIER et al. 'WashU-Merck EST Project', complete record, 13 February 1995.	1, 7-10
X	Database GenBank, US National Library of Medicine, (Bethesda, MD, USA), No. H86863, HILLIER et al. 'The WashU-Merck EST Project', complete record, 21 November 1995.	1, 7-10

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US98/27059

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(e).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

Please See Extra Sheet.

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:
1-10, 21

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.
☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US98/27059

BOX II. OBSERVATIONS WHERE UNITY OF INVENTION WAS LACKING

This ISA found multiple inventions as follows:

This application contains the following inventions or groups of inventions which are not so linked as to form a single inventive concept under PCT Rule 13.1. In order for all inventions to be searched, the appropriate additional search fees must be paid.

Groups I-XXVII, claim(s) 1-10 and 21, drawn to a polynucleotide, vector comprising same, first claimed method of use, i.e. using polynucleotide to make a cell, and the cell made by the process. Claims 1-10 and 21 recite 114 independent polynucleotides (SEQ ID NO: 11-124 or encoding SEQ ID NO: 125-238). Group I consists of the first ten polynucleotides (SEQ ID NOs 11-20 or encoding SEQ ID NOs 125-134). Each of groups II-XXVII consists of up to four of the remaining 104 polynucleotides, in order.

Groups XXVIII-CXLI, claim(s) 11, 12, 14-16 and 17 (first part), drawn to a polypeptide, a method of making the polypeptide and first claimed method of use, i.e. in treatment. These claims recite 114 independent polypeptides, each of groups XXVIII-CXLI consists of a single polypeptide as set forth in SEQ ID NOs 125-238, respectively.

Groups CXLI-CCLV, claim(s) 13 and 19, drawn to an antibody to a polypeptide and the first claimed method of using same. These claims recite 114 independent antibodies to 114 independent polypeptides, each of groups CXLI-CCLV consists an antibody against a single polypeptide as set forth in SEQ ID NOs 125-238, respectively.

Groups CCLVI-CCLXXXII, claim(s) 17(second part), drawn to an additional method of using a polynucleotide. Group CCLVI consists of methods reciting the first ten polynucleotides (SEQ ID NOs 11-20 or encoding SEQ ID NOs 125-134). Each of groups CCLVII-CCLXXXII pertains to up to four of the remaining 104 polynucleotides, in order.

Groups CCLXXXIII-CCCIX, claim(s) 18, drawn to a second additional method of using a polynucleotide. Group CCLXXXIII consists of methods reciting the first ten polynucleotides (SEQ ID NOs 11-20 or encoding SEQ ID NOs 125-134). Each of groups CCLXXXIV-CCCIX pertains to up to four of the remaining 104 polynucleotides, in order.

Groups CCCX-CDXXIII, claim(s) 20, drawn to an additional method of using the polypeptide. These claims recite 114 independent methods of using 114 independent polypeptides, each of groups CCCX-CDXXIII consists an antibody against a single polypeptide as set forth in SEQ ID NOs 125-238, respectively.

Groups CDXXIV-CDL, claim 22, drawn to a third additional method of using a polynucleotide. Group CDXXIV consists of methods reciting the first ten polynucleotides (SEQ ID NOs 11-20 or encoding SEQ ID NOs 125-134). Each of groups CDXXV-CDL pertains to up to four of the remaining 53 polynucleotides, in order.

Claim 23 is unsearchable and cannot be grouped as it is drawn to unknown and unspecified compounds.

The inventions listed as Groups I-CDL do not relate to a single inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons:

Each of the corresponding polynucleotides, polypeptides and antibodies are independent products, with different uses and being structurally, biochemically and biologically different products. Additional or alternate methods of use are claimed for individual polynucleotides and polypeptides. 37 CFR 1.475(b) does not provide for unity of invention of more than 1 product or more than one method of using a product as a combination of invention having unity of invention. However, with respect to groups drawn to independent polynucleotides or alternate methods of using same recited in the alternative, in accordance with 1192 O.G. 68 (19 November 1966) applicant is entitled to an initial search of inventions pertaining to the first ten independent polynucleotides recited, and may elect to pay an additional fee for each search of up to four additional independent polynucleotides. For additional method of using each of the independent polynucleotides, applicant may further elect to pay an additional fee for an additional search involving the first ten polynucleotides and each additional search involving up to four additional polynucleotides. With respect to groups pertaining to independent polypeptides or antibodies to the independent polypeptides, each product or method of use is an additional invention. An additional fee must be paid for search of each additional invention relating to polypeptides or antibodies against same. With respect to the relationship between the claimed polynucleotides and the claimed polypeptides, there is no one-to-one correspondence, i.e. no corresponding scope, between claims drawn to polynucleotides and their use and those drawn to polypeptides, antibodies and their use. Consequently, there is no special technical feature linking the polynucleotides and the polypeptides or antibodies claimed.